The Economic Evaluation of Youth Training Programmes

Scott W Fargher

PhD – The University of Edinburgh – 2002

Acknowledgements

My thanks to friends and colleagues at the University of Edinburgh and the University of Auckland for their support and helpful feedback on presentations of various stages of my work. In particular much is owed to my long-suffering supervisors Brian Main and David Raffe who are probably as pleased to see the finished thesis as I am.

I would also like to offer a special thank you to my friends and family for their patience and tolerance over this lengthy process and the long absences. The unceasing support and encouragement of my parents, Pat and Dick, continues to astound me but it is very much appreciated. A large dose of gratitude is also due to Steve and Lynn Jones who have been unfaltering in their encouragement and nourishment in every sense. Fabian Zuleeg too has gone far beyond the call of friendship in supporting me through this lengthy task, although he must have wondered if I was ever going to surrender his spare room. Likewise, Paul Gee who willingly shared his substantial computing knowledge and generously provided me with accommodation on my excursions through London.

I am also especially grateful for the generous financial support provided at the outset of this study by General Accident (now part of CGNU) and their support through the GA/NZI Waitangi Foundation Exchange Scholarship to Edinburgh University. In particular, Raymond Blair (at the Perth end) and Murray Lander (at the Auckland end) offered considerable encouragement and were wonderful hosts. Subsequent financial assistance came from the Department of Economics at Edinburgh University allowing me to continue and be part of a fantastic learning environment.

Finally, this exercise would not have been possible without the ongoing support, in every sense, of Marion, Dylan and Zac. I know there were lots of times you would have preferred me to remain at home – thanks guys! Finally, the thesis is my own work and any errors and omissions are down to me.

Scott, June 2002

Declaration

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person where due reference is not made in the text.

Scott W.Fargher, June 2002

Table of Contents

List of Tables and Figures

Figures

Abstract

This thesis investigates issues associated with the economic evaluation of government intervention in the youth labour market. Specifically, it focuses on the evaluation of youth training programmes and assesses the sensitivity of post-programme impact estimates to methodological choices embodied in different non-experimental methods. Relatively high and persistent youth unemployment rates, together with rapidly changing patterns in the demand for skills, demographic change and social exclusion fears, have led to the expansion of government provided education and training opportunities in most modern economies. Moreover, the increased application of 'active labour market policies' combined with the desire to maintain international competitiveness has contributed to a greater emphasis on training in most industrial countries. Clearly, the outcome of evaluation research has an important bearing on the course and scale of such intervention and is necessarily an integral part of any active labour market policy.

The use of non-experimental techniques in evaluation research has been widely questioned in light of divergent cross-study impact estimates, even when applied to seemingly similar programmes and data, together with inconsistencies when compared to estimates derived using experimental techniques and data. This thesis investigates different non-experimental methods used in economic evaluation through the detailed examination of government assisted youth training in Scotland supplemented by an examination of training provision in Australia. The experimental/non-experimental debate is used as a framework to review recent theoretical developments in the evaluation field. The analysis begins with an examination of issues in the youth labour market that have lead to increased levels of intervention. Different evaluation methodologies are then outlined followed by a detailed exploration of the institutional developments in the youth labour markets of both the United Kingdom and Australia. This is complemented by a critical review of the relevant empirical literature focussing on potential sources of cross-study variation in reported impact estimates. A meta-analysis carried out on evaluation studies in the United Kingdom confirms that a substantive part of the variation may be due to the model employed. To investigate this issue in more detail the thesis then employs sensitivity analysis to determine the impact of model choice on a common database. The empirical analysis draws on two distinct longitudinal datasets the Scottish Young Person's Survey and the Australian Youth Survey and examines initial returns to Scottish youth training in the early 1990s.

Glossary of Abbreviations

ABS	Australian Bureau of Statistics
ALMPs	Active Labour Market Programmes
ALS	Australian Longitudinal Survey
ANTA	Australian National Training Authority
AQF	Australian Qualification Framework
ATS	Australian Traineeship System
AVCTS	Australian Vocational Certificate Training Scheme
AYS	Australian Youth Survey
CES	Commonwealth Employment Service (AUST)
CETA	Comprehensive Employment and Training Act (USA)
D/EE	Department for Education and Employment (UK)
DETYA	Department for Education, Employment, Training and Youth Affairs (AUST)
DTI	Department for Trade and Industry (UK)
ENPs	Eligible Non-participants
GNVQ	General National Vocational Qualification (UK)
ITB	Industry Training Board (UK)
ILO	International Labour Organisation
ITL	Industry Training Levy (AUST)
LEC	Local Enterprise Company (SCOT)
JTPA	Job Training Partnership Act (USA)
LEC	Local Enterprise Council (SCOT)
MAATS	Modern Australian Apprenticeship and Training System
MDTA	Manpower Demonstration Training Act (USA)
MSC	Manpower Services Commission (UK)
NCDS	National Child Development Study (UK)
NCRVE	National Centre for Research in Vocational Education (USA)
NCVER	National Centre for Vocational Education Research (AUST)
NJS	National JTPA Study Demonstration (USA)
NSW	National Supported Work Demonstration (USA)
NVQ	National Vocational Qualification (UK)
NWS	New Workers Scheme (UK)
OECD	Organisation for Economic and Co-operative Development

ONS	Office for National Statistics (UK)
SYPS	Scottish Young Person's Survey
SVQ	Scottish Vocational Qualification
TAFE	Technical and Further Education (AUST)
TEC	Training and Enterprise Council (UK)
VET	Vocational Education and Training (AUST)
WIN	Work Incentive Demonstration (USA)
YCS	Youth Cohort Survey (ENGLAND & WALES)
YOP	Youth Opportunities Programme (UK)
YTS	Youth Training Scheme (UK)
YT	Youth Training

Chapter 1: Introduction

1.1: Introduction

> "Ought the public, therefore, to give no attention, it may be asked, to the education of the people? Or if it ought to give any, what are the different parts of education which it ought to attend to in the different orders of the people? And in what manner ought it to attend to them? In some cases the state of the society necessarily places the greater part of individuals in such situations as naturally form in them, without any attention of government, almost all the abilities and virtues which that state requires, or perhaps can admit of. In other cases the state of the society does not place the greater part of individuals in such situations, and some attention of government is necessary in order to prevent the almost entire corruption and degeneracy of the great body of the people" (Smith, 1976 edition, Bk. V, chap. 1, p. 302).

The youth labour market is in an almost constant state of flux with the decisions young people face in the school-to-work transition, and the opportunities open to them, having changed dramatically in recent years. The youth labour market has been characterised by high and persistent youth unemployment rates raising concerns not only about the resultant loss of production or the erosion of skills but also *social exclusion* for a large segment of the population. Such concerns together with rapidly changing patterns in the demand for skills and an ageing workforce ensure that the youth labour market remains an area of high priority in most modern economies. In the United Kingdom, for example, this is reflected in the emphasis placed by the government on policy initiatives, such as the 'New Deal', at the level of public debate, such as the need to make the market more flexible, and in the question of whether or not the education system is providing adequate skills for new entrants to the labour market. Significantly, training is frequently viewed as an imperative in maintaining international competitiveness and trading performance in the new 'knowledge based' economy.[1] Consequently, the expansion of education and

---

[1] The link between training and competitiveness has been explored in a number of recent articles, see for example Keep and Mayhew (1999) or Blundell, Dearden, Meghir and Sianesi (1999). The issue has received considerable attention in the UK resulting in a White Paper (see DTI, 1998). More detailed evidence from the UK is presented in the collection of articles edited by Stevens and Mackay (1991) on behalf of the National Economic Development Office. While a particularly persuasive argument in support of investment in skills *per se* can be found in Thurow (1999) *Building Wealth: The New Rules for Individuals, Companies and Nations in a Knowledge-based Economy*, New York: Harper Collins.

training opportunities are typically key platforms in modern government policy. Moreover, the government provision and support of training programmes and other active labour market policies for youth seem to have become a permanent feature in most economies and offer school leavers a labour market alternative to conventional employment.[2] Accurate ongoing evaluation of such interventions is imperative given the commitment of significant public resources, the presence of competing programmes, modest economic growth and an environment of increased accountability. This thesis investigates different methods used in economic evaluation through the detailed examination of government assisted youth training in Scotland.[3] The aim is to reconcile the variation in the use of non-experimental estimators observed in the existing studies conducted on training schemes in the United Kingdom. This is supplemented by an examination of training provision in Australia.

Programme evaluation is a complex area and our understanding of the economic benefits of different routes for the transition from school-to-work, to both individuals and society in general, is far from clear. For example, "to what degree does further education and training enhance the employability and remuneration prospects of young people?" It is important such issues be informed by evaluation studies employing empirical analysis. The *economic analysis* of success, in such programme evaluation, typically focuses on increased levels of employment and/or the enhancement of post-programme earnings among training programme participants. Experimental methods, which involve the random assignment of eligible participants into a non-participating control group, are often regarded as the ideal means of deriving impact estimates. Social experiments are, however, difficult to implement in such a way that can produce valid impact estimates for ongoing programmes. Moreover, they are expensive to operate and tend to be politically unpopular on ethical grounds. Thus, in practice, most programme evaluation typically employs

---

[2] That is, entering directly into employment (or unemployment) after leaving school – sometimes called the 'traditional route' in the school-to-work transition.

[3] As detailed below, the dataset I am using is Scottish, however, youth training is funded by central government in the United Kingdom and hence many of the aggregate statistics used in the descriptive analysis apply to the United Kingdom as a whole.

non-experimental methods, which utilise econometric and statistical techniques to approximate a control group. On the other hand, non-experimental analysis is not free from controversy either, as it typically necessitates complex methodologies based on subjective assumptions and consequently the results can be difficult to interpret.

Recent literature on the evaluation of manpower training programmes in the United States has demonstrated that alternate non-experimental estimators of the same, or substantially similar, programmes may produce inconsistent estimates of programme impact. Moreover, the reported impact estimates, particularly for male youth, are frequently pessimistic leading many observers to question the value of such intervention. However, such a conclusion may not be justified, as the assessment of public training programmes solely on economic impact measures are often considered to be unfair and/or incomplete. This is in part because the objectives of government provision may include elements of income transfer, contributions to public infrastructure or services, and the desire to add to the stock, or arrest the deterioration, of skills and working abilities of target groups. Equally programme objectives may be poorly defined, lack focus, overlap with the aims of other interventions and/or be subject to change. Furthermore, the evaluation process is often hindered by the nature of training programmes themselves, which have largely evolved out of employment creation programmes and typically exhibit considerable variation in both duration and quality. Nonetheless, the results of applying the same methodology to alternative training programmes (and other active labour market programmes) and judging them in terms of the same outcomes can provide policy-makers both with information about the implicit trade-offs involved when choices must be made between programmes and an effective means of monitoring programme development.

Ongoing government commitment together with changes in the design and delivery of training programmes further increases the pressure for more accurate and detailed evaluation of government assisted training programmes. Utilising rich national longitudinal data from Scotland, this thesis examines the behaviour and applicability of different non-experimental methods used in the estimation of returns from youth training programmes. This issue is of significant practical importance for several reasons. First, given the fiscal cost of such policies, it is necessary to assess rigorously whether this spending has achieved its objectives, however defined. Second, existing studies have reported a diverse range of estimates, even when applied to substantially similar programmes employing the same or similar dataset. Third, randomised experiments cannot always be implemented due to practical and ethical reasons and the vast majority of data available in the social sciences remains non-experimental. Fourth, comparing the relative efficiency of different estimators using a common database may aid the choice of the appropriate estimator. In addition, even though the evaluation of youth training schemes merits substantive research in its own right, it can also serve as a prototype for deriving impact estimates in a variety of closely related fields ranging from unionism to migration.

A detailed critique of the relevant theoretical and empirical evaluation literature precedes the empirical analysis. This approach is designed to identify possible sources of variation observed in reported estimates of training impact.[4] In some cases it is then possible to control for the factors contributing to the variation in the subsequent analysis and focus on the sensitivity of estimates to different model constructs. The aim is to enhance our understanding of the relative behaviour of different non-experimental estimators, make any trade-offs explicit and contribute to more accurate estimates of training impact. The next section presents a more detailed overview of the remaining chapters in this thesis.

1.2: Thesis Overview

The context for evaluation research is provided in chapter 2, which addresses contemporary issues in the youth labour market focusing on the United Kingdom (UK) and Australia. This is achieved by developing a statistical profile for each of the respective youth labour markets. It is shown that, in general, the state of the youth labour market falls short of what is desirable with youth unemployment rates

---

[4] As already observed, the existing evaluation literature has produced a diverse range of reported estimates for training impact. It is this inconsistency in impact estimates that has provided much of the motivation for undertaking this thesis.

remaining relatively high and persistent in both countries. Alarmingly, youth unemployment rates have remained high relative to aggregate unemployment rates despite the fact that significant proportions of teenagers are choosing to stay on at school and more are leaving school with qualifications. This trend is particularly prevalent for young men and, in the case of Australia, amongst indigenous youth. Moreover, drawing on OECD data, the chapter illustrates that this pattern is reflected in a wide range of advanced industrial countries. Chapter 2 also examines trends in terms of government intervention highlighting the increased role of active labour market policies such as youth training. Different evaluation issues in economics are introduced including the key 'evaluation problem' of determining an effective counterfactual in the presence of *sample selection bias*. This form of bias is pervasive throughout much of the social sciences empirical literature, and arises because one cannot simultaneously observe the same person both in a programme and out of it. Additionally, a number of key definitions as used in the subsequent analysis are clarified.

Chapter 3 then addresses evaluation issues beginning with an outline of the evaluation problem in intuitive terms before undertaking a detailed examination of the experimental versus non-experimental debate that has underpinned recent evaluation literature. This approach provides us with a ready vehicle for reviewing the major theoretical developments in the relevant literature. It is also clear that the non-experimental approach can benefit from recent advances made in the experimental field, particularly the need to establish and maintain the validity of the control group and to consider the estimation of programme impacts other than the mean. Significantly, limitations on the application of social experiments, on practical and ethical grounds, together with the much wider availability of non-experimental data necessitates the continued application of non-experimental techniques. Moreover, as observed by Ham and LaLonde, (1995, p. 231) "…even in experimental settings, we must often rely on non-experimental methods when we turn from examining whether training works to how it works". It is hence important to improve our understanding of how non-experimental methods work and to identify what trade-offs are involved in applying different methods. Chapter 3 also discusses extensions to the basic sample selection model and outlines different methods for evaluating alternative non-experimental estimators. The focus for the remainder of the thesis is primarily on the non-experimental approach and the choice of an appropriate estimator.

Chapter 4 presents an overview of recent developments in the institutional arrangements for the provision of youth training in both the UK and Australia before reviewing the empirical literature on youth training programmes in these countries. This illustrates the array of different policies adopted in each of the countries, which have the aim of addressing the substantial disequilibrium that has beset their youth labour markets in recent decades. The main focus of the review of evaluations is on government supported youth training programmes in the UK due to the greater availability of relevant studies and articles. The reported estimates clearly illustrate the wide range in estimates of programme impacts inherent in non-experimental evaluations of similar programmes using seemingly similar datasets. The UK studies are compared and contrasted with programme evaluations in the USA drawing on Barnow's extensive 1987 review of non-experimental studies. A number of more recent studies from the USA are also added in order to provide some comparison with experimental estimates. There have been comparatively few evaluations made on comparable government supported programmes in Australia where the bulk of evaluation work has focused on the provision of training in the private sector, for example see the extensive work of Dearden (1995). In this section, a selection of existing Australian studies, which largely rely on aggregate data, are reviewed and related studies are discussed. This chapter outlines what we can learn from these earlier studies and identifies possible sources of variance in the reported estimates by providing a methodological classification of the studies reviewed in the chapter.

In Chapter 5 the results of a meta-analysis on the UK studies is reported in an attempt to more objectively identify the sources of variation in estimates across these studies. A meta-analysis is a quantitative review of reported literature findings using the summary statistics from the original studies as the data-points for analysis. The aim is to achieve greater accuracy than that obtainable from a single study and also to allow the investigation of factors responsible for between-study variation. Chapter 5 includes an overview of the development of meta-analysis and review of existing meta-analyses in economics. The results of the analysis suggest that model choice contribute significantly to the observed variation in estimates in the UK evaluation literature. On the other hand, the direction of impact identified in the analysis for the different outcome measures is consistent with the conclusions drawn in the previous chapter and in other traditional literature reviews.

The next chapter, chapters 6, reports the returns to government-supported youth training programmes in Scotland, both in terms of the impact on earnings and the probability of subsequent employment. In this chapter data from the Scottish Young Person's Survey (SYPS) is used to estimate the returns to youth training in the early 1990s for a cohort of young people who reached the minimum school leaving age (16) in October 1989. The chapter includes a detailed description of the datasets along with an overview of relevant institutional factors and prevailing economic conditions in order to establish the context of the evaluation. A sensitivity analysis is conducted through the application of different non-experimental methods and specifications to systematically determine any sources of variation in earnings estimates that may be due to model choice. The aim is to demonstrate how the various assumptions inherent in the application of different evaluation methods can influence the derived impact estimates. Earlier studies, in seeking to explain the observed variation in estimates, raised a number of different factors that should be taken into account by analysts but they evaluated different programmes and/or utilised different datasets. Consequently, these attempts typically only addressed a sub-set of possible economic and statistical misspecifications and the insufficient overlap of the different studies did not allow any firm conclusions relating to the relevance of these factors to be drawn. In contrast, the application of sensitivity analysis accounts for any variance that may arise from the analysis from a range of exogenous factors, such as different programmes and or datasets. The sensitivity analysis includes the choice and development of a baseline specification, the application of relevant specification testing procedures and employs the main methods used to generate estimates in earlier empirical studies. The analysis illustrates that much of the variation in UK estimates of training programme impact is likely to be due to the use of alternative evaluation methods. The fact that the various models arrive at different impact estimates indicates that sample selection bias is inherent in the dataset and raises a number of issues related to the use of this type of survey data for evaluation purposes.

A second dataset, the Australian Youth Survey (AYS), was examined with the intention of applying similar specifications to evaluate government supported youth training in Australia. However, given the different structure of the dataset and limited observations detailing the receipt of government assisted training combined with the nature of support for youth training in Australia, which is contractually based and delivered through a range of schemes, this was not feasible. Given the extensive overlap in public and private sector provision consideration was given to evaluating returns to training in general but the AYS has already been used to evaluate training in this fashion in a doctoral thesis presented by Lorraine Dearden (1995).[5] The description of training provision in Australia proves useful in the sense that it identifies a number of similarities in recent developments with the provision of training in the UK and, hence with the availability of suitable data, many of the techniques applied in chapter 6 would be readily transferable to the Australian setting.

Finally, chapter 7 concludes the thesis, summarises the arguments of previous chapters, identifies potential applications of this research and maps out directions for future research. It also includes possible policy implications arising from the analysis and a discussion on the design and collection of non-experimental data necessary for more effective training evaluation.

---

[5] A description of the AYS dataset is given in the appendix to this thesis.

Chapter 2: Background Issues in the Youth Labour Market

2.1: Introduction

Young people face a variety of unique challenges on entering the labour market, ranging from occupational choice to having insufficient or mismatched skills. This chapter examines the different characteristics of the youth labour market that have lead to substantial state intervention, such as the provision of government assisted training programmes. A statistical profile is developed for the youth labour markets in both Australia and the United Kingdom supported by aggregates from other industrial nations to illustrate the extent of the challenge facing policymakers. From a policy perspective, intervention in the market is typically designed to help address the immediate problem of youth unemployment and the longer-term concerns of a possible deficit in the skills on which an economy depends. Policy may also be directed to assisting those considered 'at risk' or disadvantaged in the labour market, such as ethnic minorities, unqualified school leavers or solo parents. The scale of such intervention and the presence of competing programme options together with an environment of increased accountability have impacted on the demand for accurate ongoing evaluation of these programmes and hence provides the context for this thesis. Moreover, in recent decades the youth labour markets in most industrial nations have tended to exhibit relatively high and persistent unemployment rates, a high incidence of long-term unemployment and rapidly changing skill demands on new entrants. These factors have combined to ensure that the youth labour market remains an area of high priority in most modern economies. Government supported youth training programmes have been one of the pivotal responses to the challenge of youth unemployment and skill formation as part of a general trend toward increased use of 'active' labour market programmes.

In section 2.2 the distinctive characteristics of the youth labour market are examined in terms of unemployment and activity measures allowing assessment of the perception that there are inherent shortcomings in the market justifying intervention. This is further investigated by examining these measures relative to the corresponding adult rates. Section 2.2 also discusses demographic change and the important issue of long-term youth unemployment. Section 2.3 discusses the international trend towards increased emphasis on and application of active labour market policy. This section also provides an overview of the main labour market policy categories, with particular reference to youth measures. A brief discussion of recent youth policy developments in the United Kingdom and Australia is also included in section 2.3, however these are examined in more detail in chapter 4. The evaluation problem, which arises from our inability to observe an individual in more than one state at any point in time, is then introduced and the different approaches economists use for the evaluation of labour market interventions are outlined in section 2.4. A summary of the chapter is then presented in section 2.5.

2.2: Characteristics of the Youth Labour Market

The youth labour market presents the researcher with a number of unique challenges, not least amongst these is determining the relevant age group for analysis. The youth labour market is generally defined as comprising those people in the fifteen to twenty four years age group, (inclusive) although the lower limit usually depends on the statutory minimum school leaving age. In the United Kingdom (UK) and the United States of America (USA), for example, youth policy is directed to the 16 to 24 year old age group. This definition is typically used as the standard when reporting youth labour market statistics but tends to be more loosely applied in matters of policy when the upper limit, in particular, is subject to wide variation. Until recently, youth employment policy in the UK was targeted at the 16 to 18 year old age group, however, it has been widened to include those people aged up to 24 years with the introduction of the New Deal.[6] In contrast, Italy's youth policy generally targets the 14 to 29-year old age group in Northern Italy and the 14 to 32 year age group in the South (O'Higgins, 1997, p. 3).

The concept of a youth labour market is further blurred by changes in the school-to-work transition due to educational and training reforms, which have been designed to

---

[6] More recently the New Deal has been further extended to include those aged 25 and over however, for this age group, the programme concentrates on assisting those who have been unemployed for two years or more.

encourage young people to spend longer in education or even combine education and work. A number of countries have a *dual* system, formally combining education and work, such as the well-established dual apprenticeship schemes in Austria, Germany and Switzerland. In the 1980s and 1990s other countries introduced a proliferation of vocational training programmes, targeted at the young, which contain both work and study elements, such as those in France, the UK and Sweden. In other countries a clear distinction remains between school and work, such as in Italy, Spain and to a great extent Australia (OECD, 1998, pp. 234-241).[7] In addition, many countries have withdrawn or restricted social security benefits to young people. These changes, combined with an uncertain job market, appear to have contributed towards making the school-to-work transition both more difficult and more prolonged for many young people.

Unless otherwise stated, the standard definition for the youth labour market (15 to 24) is employed throughout this thesis. Where applicable, however, the distinction between teenagers, the 15 to 19 year old age group, and young adults, the 20 to 24 year old age group is drawn as these groups often face distinct challenges.

In economic terms, examination of just three key statistics can help us assess the state of the youth labour market: the unemployment rate, the labour force participation rate, and the employment/population ratio. Although closely related, these statistics can sometimes be at variance with each other. For example, Clark and Summers (1982) showed that teenage unemployment rates often remain fairly high even when aggregate unemployment declines due to increased participation encouraged by the economic upswing. Nevertheless, these three measures can be used to establish a broad overview of trends in youth labour market activity. The remainder of this section utilises data from the Organisation for Economic Co-operation and Development (OECD) for this purpose. Additional detail is provided

[7] Raffe, Biggart, Fairgrieve, Howieson, Rodger and Burnisten (1998) describe the British education and training systems as 'notoriously difficult' to summarise in a simple description or model. This is largely due to flexibility inherent in the design and delivery of post-compulsory pathways and qualifications. Similarly, recent reforms in Australia encourage greater overlap between education and work. Further discussion on these issues can be found in chapter 3 and 4.

on our countries of interest, Australia and the UK,[8] including the duration of unemployment and relevant demographic factors.

2.2.1: Unemployment

The unemployment rate is the most commonly employed indicator of youth labour market conditions and the generally accepted standard is the definition applied by the International Labour Organisation (ILO). The ILO defines the unemployed as those people who have not worked more than one hour during the reference period but who are available and actively seeking work (O'Higgins, 1997, p. 3). Another common measure is claimant unemployment, referring to only those registered for unemployment assistance. National measures, employed by individual countries, also tend to vary significantly. Most countries, for example, exclude students, defining them as being outside the labour force but some countries, such as Norway, include them if they are actively seeking work. Some countries interpret the one-hour rule more liberally; the Netherlands for example, classifies anybody working less than twelve hours per week as unemployed. Moreover, the standard measure does not include discouraged workers, those not employed but not actively seeking work, which may make up a large portion of youth particularly when the economy is in recession (Stern, Bailey and Merritt, 1996, p. 7). Different measures obviously limit comparative analysis and consequently, for the purpose of this thesis, the ILO definition is employed unless otherwise stated.

Throughout most modern economies the youth labour market has tended to be characterised by acute and persistent unemployment in recent decades. Youth rates are invariably higher than adult rates and in many countries they appear to be on an increasing trend. In some countries this may be explained by the movement into recession, however in many countries it appears to be part of an underlying structural

---

[8] England, Wales and Scotland each has its own education and training system, however, all three systems share 'British' features. Significantly, Raffe et al. (1998) note that "...their current differences may be less important than the potential for future divergence (or possibly convergence), especially following the establishment in 1999 of the National Assembly in Wales and the Parliament in Scotland" (p. 5). Where possible Scottish statistics are reported alongside UK aggregates.

problem. Between 1970 and 1994, for example, youth unemployment rates in the 11 main European countries rose, on average 16 percentage points, from 4.2 percent to 20.2 percent (Korenman and Neumark, 1996). In contrast, youth unemployment rates rose from 2.3 percent to 16.2 percent in Australia and from 2.6 to 17.2 percent in the UK, over the same time period.[9]

Figures 2.1a and 2.1b, below, utilise OECD data to illustrate differences in relative unemployment rates by age group for men and women in the UK, between 1988 and 2000. For both sexes the youth rates are nearly twice the level of those for adults and rarely fall below 10 percent. This virtually replicates the pattern from the previous decade. Examination of OECD data suggests that this pattern is common to most advanced industrial countries, including Australia. This is illustrated in Table A1, in the appendix to this thesis, which lists the youth and adult unemployment rates in 1988 and 1998 for 16 OECD countries. Columns 1 and 2 of the table illustrate that, to a greater or lesser extent, the pattern displayed in tables 2.1a and 2.1b holds in the majority of countries listed. In 1988 the average youth unemployment rate in OECD member countries was 14.4 percent compared to only 6 percent for adults, by 1998 the respective averages were 13 and 5.7 percent.

It is also evident from the series illustrated in figures 2.1a and 2.1b that there is considerable variation within the youth age group with unemployment being much more prevalent amongst teenagers (16 to 19 years) than for young adults (20 to 24 years). This is particularly noticeable for young women in the UK (figure 2.1b). The variation may indicate more difficult market conditions, lower qualifications, and inexperience or, more simply, may be a reflection of changing participation rates for this group (see section 2.2.2 below).

[9] In contrast, over the same period, youth unemployment rates in the USA moved only 1.5 percentage points. Starting from a much higher base of 11 percent in 1970, youth rates rose to 12.5 percent in 1994. This substantially improved the USA's position relative to other OECD countries. To a large degree however, this reflects the robustness of their economy in 1994 compared to other industrialised countries rather than a healthier youth labour market *per se*.

Figure 2.1a: UK unemployment rates by age, 1988 to 2000 - Males*

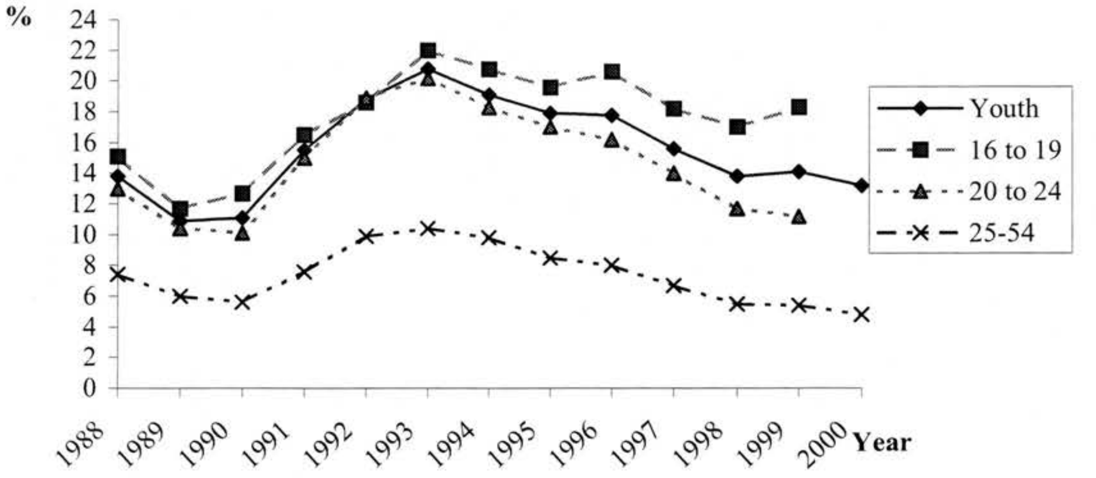
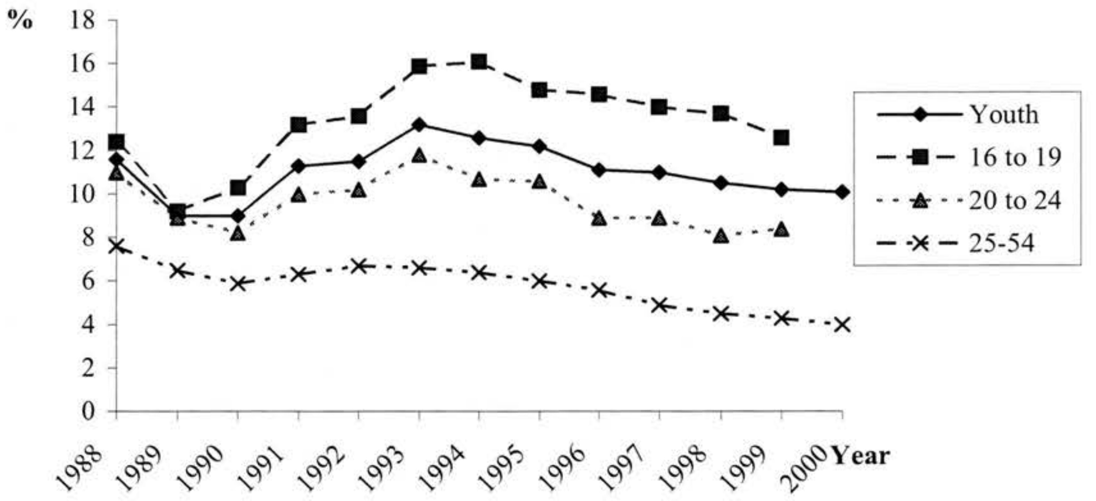


Figure 2.1b: UK unemployment rates by age, 1988 to 2000 - Females*



Source: 1988-99 figures – OECD Labour Force Statistics, 1997-1999
2000 figures – OECD Employment Outlook, 2001 June

** Youth rate refers to 16 to 24 year age group*

Further inspection of figures 2.1a and 2.1b suggests that variations in the youth rate appear to be directly proportional to variations in the adult rate although more sensitive to cyclical fluctuations. In a recent cross-country investigation of this relationship Gaude (1997) estimated that the elasticity of the youth unemployment rate with respect to the adult rate was close to one. This indicates that, given the higher relative youth rates in most countries, youth unemployment is more sensitive to fluctuations in economic conditions in absolute terms. Another measure that captures the relative performance of the youth labour market is the youth/adult unemployment ratio measure. In the UK the ratio stood at 1.96 in 1986 but by 1996 it had risen to 2.18 in spite of corresponding improvements in the absolute measures of unemployment. In other words, youth unemployment was falling at a slower rate than adult unemployment over this period (Raffe et al. 1998, pp. 4-5). In most OECD countries however, the inverse is observed. The pattern in Australia where the ratio has declined from 2.8 to 2.1 over the same time period is more typical (see Ainley, Malley and Lamb, 1997).

Youth unemployment measures are also subject to variation by gender as illustrated in figure 2.2a and 2.2b. The youth rate for males consistently exceeded the female rate in the UK between 1988 and 2000. This pattern holds in the UK for both teenagers and young adults. Figure 2.2b shows that the male youth unemployment rate in Australia also tended to exceed the female rate for much of this period. Examination of the disaggregated series, however, reveals a more mixed pattern. Unemployment rates for female teenagers tend to exceed those of their male counterparts while the inverse is observed for young adults.

Figure 2.2a: UK youth unemployment rates by gender, 1988 to 2000[1]

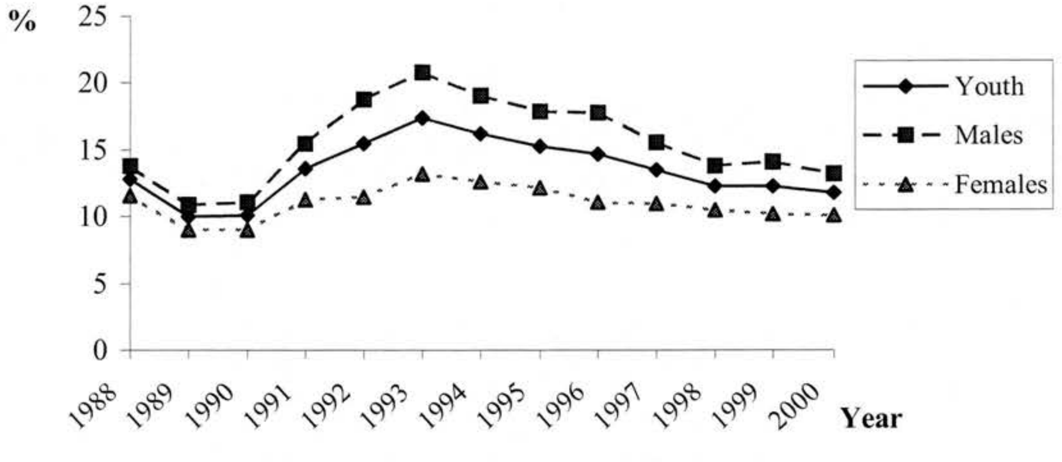
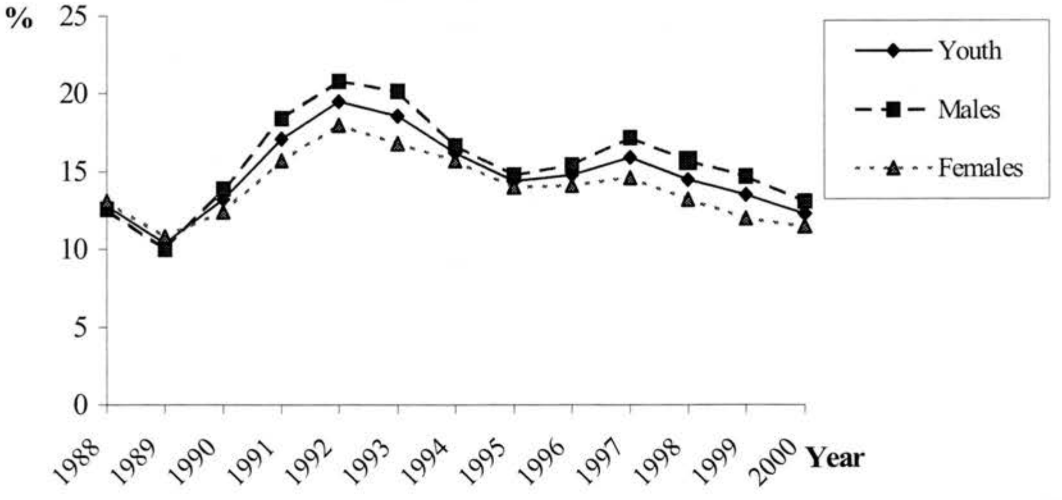


Figure 2.2b: Australian youth unemployment rates by gender, 1988 to 2000[2]



Source: 1988-97 figures – OECD Labour Force Statistics, 1977-1999
2000 figures – OECD Employment Outlook, 2001 June

[1.] *Youth rate refers to 16 to 24 year age group*
[2.] *Youth rate refers to 15 to 24 year age group*

In Australia considerable variance is also observed between Indigenous and non-Indigenous youth unemployment rates. A recent study by the Department for Education, Employment and Youth Affairs (DETYA)[10] reported that youth of Aboriginal and Torres Straits Islander descent are just under twice as likely to be unemployed as non-indigenous youth. Rural unemployment rates are substantially higher for both groups but the same pattern of higher indigenous unemployment is observed.[11] Improving the employment prospects of Australia's indigenous youth is thus a key priority for the DETYA.

There are a number of factors that may explain the relatively high youth unemployment rates. These range from problems in gaining a first job following the completion of schooling due to insufficient skill or experience, to a greater propensity to be laid off during swings into recession. The latter arises from the often applied practice of 'last on first off' as well as the tendency toward temporary contracts observed in many OECD countries which lowers the opportunity cost of firing young people (see OECD, *Economic Outlook, 1996 June,* chapter 1). A related factor is that many firms respond to recession by ceasing to hire new workers, before embarking on the more expensive procedure of redundancies. This would have a disproportionate impact on youth, as they comprise a considerable proportion of new job seekers (O'Higgins, 1997, p. 7).

There is also evidence in some countries that young people are more likely than their older counterparts to voluntarily quit their jobs as part of a process of shopping

---

[10] The DEETYA was formally the Department of Education, Employment and Training (DEET) and more recently has dropped Employment from its title adopting the shortened acronym DETYA. For the purposes of uniformity the acronym DETYA will be used throughout the thesis. Similarly, recent changes in the UK have seen most of the responsibilities previously vested in the Department for Education and Employment (D*f*EE) now come under the auspices of the newly formed Department for Education and Skills (D*f*ES). Some of its responsibilities have been assigned to the Department for Work and Pensions (D*f*WP). However, for the sake of uniformity the acronym D*f*EE will continue to be used throughout the thesis (currently the D*f*EE web address links one directly to the D*f*ES website).

[11] The Department's report, *The School to Work Transition of Indigenous Australians,* used data from the 1991 and 1996 census and is discussed in more detail in section 3.6 of this thesis. Indigenous Australians make up 2.1% of the total population and 2.6% of the population of 15 to 19 year olds. This means that Census and administrative collections provide the relatively few reliable sources of information.

around for a more productive job match (Buechtemann, Schupp and Soloff, 1993). For many young people, high job mobility or 'churning' is the mechanism by which they progress towards a more permanent job, a pattern that might characterise the United States, Canada, Australia, and to a lesser degree, the United Kingdom. The dual systems in Europe, however, appear to move young people into stable employment more smoothly involving fewer jobs and less time (OECD, *Employment Outlook*, 1993, 1996). Analysis typically focuses on the length of time or number of jobs between entering the labour market and finding a stable or 'good' job although the direct relationship between churning and higher relative youth unemployment is rarely addressed (for example, see Dolton, Makepeace and Treble, 1994c or OECD, *Employment Outlook*, 1996, June). Moreover, there is no clear consensus as to whether such churning, where it is present, is reflective of some systemic failure in the youth labour market or whether it allows young people to pick-up a more diverse, and necessary, set of employment experiences, that is a more healthy youth labour market (Lynch, 1993; Barwell, 2000).

Another pervasive argument is that apparent higher relative youth wages, with respect to those of adults, have a negative effect on youth employment. The debate surrounding the recent introduction of a minimum wage in the UK centres on similar arguments.[12] This argument, however, rests on the assumption that adult and youth workers are close substitutes, which may not hold with respect to demand for skilled workers (O'Higgins, 1997, p. 7).[13] Higher youth wages are more likely to induce employers to substitute lower productivity workers with higher productivity workers which, in turn, will be encouraged to substitute remaining in school for work

[12] A minimum wage of £3.20 per hour was introduced in April 1999, accompanied by a lower rate of £3.00 applying to 18-20 year olds (and initially to 21 year olds). However, it does not apply to 16 and 17 year olds, formal apprenticeships, or to workers starting a new job with an employer and receiving accredited training. The rate was increased to £3.70 in October 2000 but the impact of the introduction of a minimum wage, in terms of employment, can only be assessed over time. It may also impact on the provision of training as argued in a recent paper by Neumark and Wascher (1998) who examined variations in minimum wage laws in the USA between 1981 and 1991. They found that the minimum wage reduced employer provided on-the-job training, particularly for the 20 to 24 years age group.

[13] In both our countries of interest the level of educational attainment is slowly rising, but it is particularly evident in the 20-24 years age group. For example, the proportion of young adults holding a senior secondary or post-school qualification in Australia rose from 68 percent in 1991 to 79 percent in 1998 (DETYA, Annual Report, 1999).

(Neumark and Wascher, 1994). In a recent study Blanchflower and Freeman (1999, p. 17) observe that, for most OECD countries, youth workers have in fact experienced declining earnings relative to adult workers through the 1990s but there is little evidence that this has had any beneficial impact on the youth labour market.

Other factors in the makeup of the youth labour market also impact on the unemployment rate, primarily the changing patterns in labour market activity. Labour market activity is itself partly determined by the state of the labour market and the availability of viable alternatives to conventional employment (and unemployment), such as government assisted training or job placement schemes. The labour market participation rate and the employment/population ratio are the primary measures used for monitoring labour market activity.

2.2.2: Participation and Employment

Labour market participation rates and the employment/population ratios have tended to fall for young people throughout the 1980s and 1990s in most modern economies. In general, the two measures closely follow each other. The participation rate provides a ready overview of labour market activity but like unemployment excludes certain categories such as discouraged workers. The employment/population ratio (employment rate), which includes the entire population in the denominator, makes no attempt to separate those actively seeking work from those who aren't, although this does not necessarily make it a more valid statistic (Stern et al., 1996, p. 7).[14]

The international trend of declining participation and employment rates is again illustrated in table A1, in the appendix to this thesis, which lists the participation rates and employment /population ratios for youth and adults in 16 OECD countries for 1988 and 1998. Columns 3 to 6 of the table show that, on average, youth participation rates fell in this period from 57.4 to 51.7 percent and the employment/population ratio for youth declined from 49.4 to 45.1 percent. The trend

---

[14] Stern, et al. (1996) note that none of the employment or unemployment measures say anything about the quality of jobs that people hold. Clearly, this observation also holds when employment is the outcome of interest in the evaluation of training programmes.

can however be seen more clearly below in table 2.1, which presents a breakdown of the relevant rates for the United Kingdom and Australia, by age group and gender.

The trend to lower youth labour market participation is most clearly seen in male rates, given for 1979, 1989 and 1999 in the upper portion of the table. In Australia, for example, the participation rate of male youth (15 to 24 years olds) was 75.3 percent in 1979 falling to 70.8 percent in 1999. This is mainly due to the pronounced drop in the labour market participation of teenage males (15 to 19 years), which had declined from 61.4 percent in 1979 to just 54.3 percent by 1999. Adult male participation rates (25 to 54 years) also declined over this period but remained at 90 percent or more. The picture in the United Kingdom is more mixed with the majority of the different age groups recording increased participation rates between 1979 and 1989 before declining in the 1990s. The participation rate of male youth (16 to 24 years) was 83.7 percent in 1989 falling to 73.2 percent in 1999. Again this is mainly due to the pronounced drop in the labour market participation of teenage males (16 to 19 years), which declined from 74.5 in 1988 to just 64.1 percent in 1999. In contrast, adult rates remained relatively stable over the same period with the respective participation rates at 94.9 and 91.6 percent.

A similar but less pronounced pattern is observed in the youth employment/ population ratio for males, presented in the lower portion of table 2.1. For male youth in Australia the ratio fell from 67 percent in 1979 to 60 percent in 1999 with the teenage ratio declining from 52.5 to 42.2 percent over the same period. For males in the UK, the ratio was only available for 1989 and 1999 but the same patterns are clearly evident with the teenage effect again dominating.

Female participation and employment rates in Australia and the UK have been at similar levels to those of their male counterparts for much of the 1980s and 1990s, particularly for teenagers. The uniformity begins to diverge, however, as they move into young adulthood with relatively more women then men out of the labour force, presumably for reasons associated with child-bearing and child-rearing.[15] The table also clearly illustrates the trend toward increased labour market activity amongst adult women in both countries.

Table 2.1: Youth/Adult labour market statistics in Australia and the United Kingdom, 1979, 1989 and 1999

a) Labour Force Participation Rate (%)

Country \ Age group			*15-19*	*20-24*	*Youth (15-24)*	*Adults (25-54)*
Australia	Males:	1979	61.4	90.2	75.3	94.5
		1989	59.5	89.2	72.6	92.7
		1999	54.3	87.1	70.8	90.0
	Females:	1979	55.0	69.2	61.9	51.4
		1989	57.1	77.5	65.6	63.0
		1999	56.2	75.5	65.9	69.2
UK[a]	Males:	1979	71.6	85.9	79.2	95.7
		1989	74.5	91.2	83.7	94.9
		1999	64.1	81.2	73.2	91.6
	Females:	1979	70.4	67.0	68.6	63.3
		1989	70.1	72.7	71.6	70.6
		1999	59.7	69.8	65.0	75.9

b) Employment/Population Ratio

Country \ Age group			*15-19*	*20-24*	*Youth (15-24)*	*Adults (25-54)*
Australia	Males:	1979	52.5	82.6	67.0	91.7
		1989	51.9	82.1	62.8	89.0
		1999	42.2	71.3	60.3	85.0
	Females:	1979	43.8	63.6	53.5	48.8
		1989	48.8	71.4	59.3	62.3
		1999	43.7	66.2	58.0	65.6
UK[a b]	Males:	1989	65.8	81.6	72.6	89.2
		1999	52.1	71.6	62.9	86.7
	Females:	1989	64.3	69.2	65.9	67.3
		1999	52.4	63.9	58.4	72.6

Source: OECD Labour Force Statistics, 1977-1999
OECD Employment Outlook, 2001 June
[a.] *Youth measures are for 16 to 24 years of age and teenagers 16 to 19 years. All the measures for the United Kingdom exclude Northern Ireland.*
[b.] *1979 figures not available*

[15] Detailed discussion on activity measures in the youth labour market for a wide range of OECD countries between the 1970s and early 1990s can be found in Blanchflower and Freeman (1996 and 1999).

To a large extent the fall in participation and employment rates has been associated with higher *staying-on* rates in education. The OECD has estimated that the average number of years a person aged 15, in a member country, could expect to be in education was 5.3 years between the ages of 15 and 29 in 1985 increasing to 6.7 years in 1996 (OECD, 1988, pp. 243-248). In England the increased education participation in the 1990s has been highly significant, with the percentage of 16-year-olds still in full-time education growing from 51.5 percent at the end of 1988 to 70.5 percent by the end of 1998. The relative percentages for 17 year olds were 35.5 and 58.7. While the percentage of 18 year olds choosing to remain in full-time education approximately doubled over this period, from 18.7 to 37.3 percent. Taken together with training schemes the percentage of 18 year olds remaining outside the labour market in 1998 stood at nearly 60 percent, compared with approximately 40 percent at the end of 1988 (DfEE, *Statistical First Release*, June 1999).

These changes have largely been contributed to the introduction of a new examination in 1986, the General Certificate in Secondary Education (GCSE),[16] which replaced the two sets of examinations formally taken at the end of compulsory schooling, at age 16 (Heath and Cheung, 1999; Robinson, 1996). The first full cohort sat the GCSE in 1988 and this is closely linked to the observed higher staying on rates. Indeed Robinson (1996, p. 79) goes so far as to contend that the:

> "...GCSE is probably the single most important reason for the increase in enrolment in further education as it led to improved attainment at age 16".

The significance of this, given the context of this thesis, lies in the fact that the level of attainment in these examinations, which has continued to improve annually since 1986, is the main determinant of destinations at the end of compulsory schooling (Raffe et al. 1998, p. 26). This is illustrated in table 2.2, which gives the status of a cohort of youth in England, by qualification in terms of GCSE grades in 1996. Young people who attain higher grades are far more likely to remain in full-time education while those who gain lower qualifications are more likely to enter the

---

[16] The equivalent qualification in Scotland is the Standard grade. English and Welsh students typically take up to eight or nine subjects for GCSE while Scottish students take up to seven or eight Standard grades and each subject is graded on a seven-point scale, from A* [highest] to G [lowest] (Raffe et al. 1998, p. 32).

labour market or youth training. Significantly, young people who fail to gain any qualifications are at much higher risk of unemployment.

Table 2.2: Post compulsory schooling status by qualification in the United Kingdom, 16-19 year olds[1]

Status / GCSE awards	*Percentage of Cohort*	*Full-time education*	*Youth training*	*Employed & training*	*Employed not training*	*Other*
5+ at A* - C	45	58	11	22	12	7
1-4 at A* - C	27	26	32	33	34	25
5+ at D – G	18	12	38	25	31	29
1-4 at D – G	4	2	7	8	7	14
none	6	2	11	12	17	26

Source: Adapted from table 3.4.2 in Raffe et al. (1998, p. 37).
[1.] Pertains to 1996, YCS data.
The 'Other' category includes those not in employment, education or training.

These trends are also observed in Scotland where, using four successive cohorts of the Scottish Young Person's Survey (SYPS), Patterson and Raffe (1995) found that in the period 1984/85 to 1990/91 fewer young people were going straight into work or training and more were remaining in education. The percentage of 16 year olds still in full-time education grew from 48 to 63 percent over this period. Scottish Office figures confirm that this trend is continuing with 66.7 percent of sixteen year olds remaining in full-time education in the 1996/97 school-year (Scottish Office, 1999).

In Australia, the increase in educational participation rates has been even more dramatic where the proportion of teenagers staying-on at school has jumped from one-third to just under three-quarters (74 percent) in the decade ending 1996 with approximately two thirds completing year 12 (senior secondary school). This increase coincided with the introduction of an education allowance, AUSTUDY, in 1987.[17] AUSTUDY is payable directly to 16 and 17 year olds remaining at school, but subject to means testing on their parents' income. Dearden and Heath (1996)

[17] The maximum AUSTUDY allowance payable to a 17-year-old is roughly equivalent to the Youth Training (YT) allowance paid in the UK.

evaluated the impact of AUSTUDY on school retention rates of children from families that are relatively disadvantaged. They identified a modest positive effect attributable to AUSTUDY (approximately 3.5 percent between 1987 and 1993) but observed a significant upward trend in participation rates, which appears unrelated to the AUSTUDY policy reforms (pp. 26-27). Retention rates have subsequently declined partly explained by economic recovery in Australia and the increased availability of full-time jobs for youth.

Amongst Indigenous youth the proportion staying-on until year 12 remains significantly lower than for non-Indigenous youth as illustrated in table 2.3. A notable improvement has however been witnessed within the last decade with the retention rate rising from 25 percent in 1992 to 35 percent in 1999. Significantly, for our purposes, a sizeable proportion (approximately one-quarter) of early school-leavers in both groups subsequently participate in vocational training. Department figures also indicate that young people of low socio-economic status and/or from rural and remote districts are also more likely to be amongst early school-leavers. On the other hand, females are more likely to remain in school with the retention rate having exceeded that for males since the mid-1970s, standing at a difference of 12 percentage points in 1999 (DETYA, *Annual Report*, 2000).

Table 2.3: Retention rates in full-time education for Indigenous Australians, 1989-1999 (%)

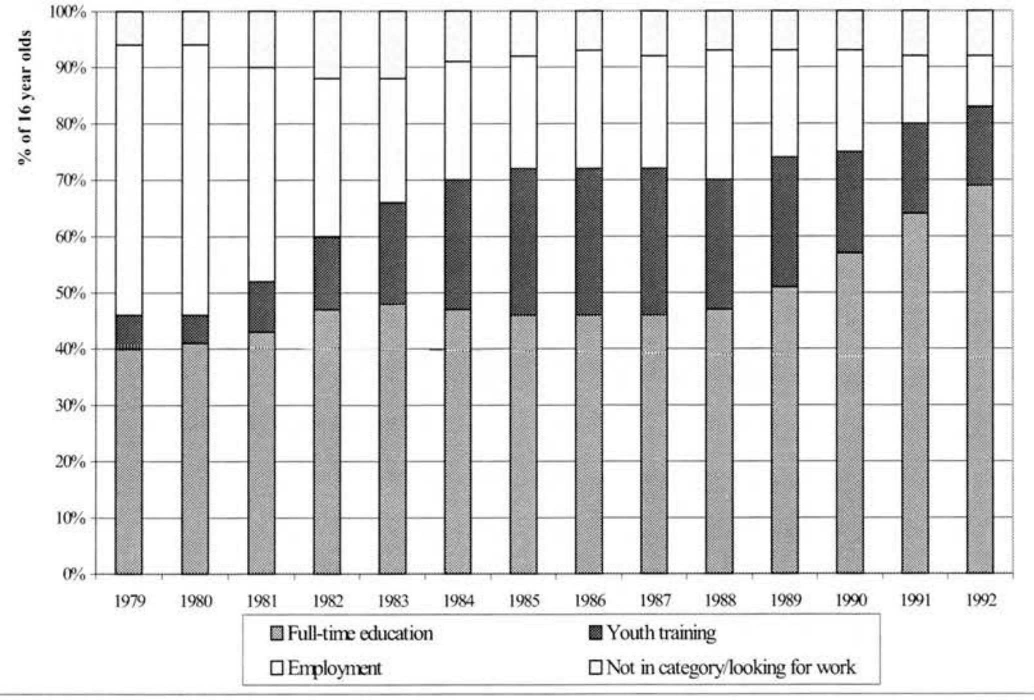
	1989	*1990*	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*
Indigenous Students											
Year 10	72	75	82	86	81	79	76	76	81	83	82
Year 12	n.a	n.a	n.a	25	25	33	31	29	31	32	35
All Students											
Year 12	60	64	71	77	77	75	76	74	73	72	72
Estimated completion rate											
	n.a	n.a	69	69	69	68	67	65	64	66	67

Source: DETYA Annual Report (various years)
n.a: figures not available

The conscious decision by many governments to increase educational participation has undoubtedly had a significant influence on the upward trend in staying-on rates however continuing enrolment beyond the minimum statutory school leaving age also reflects the state of the youth labour market. Poor labour market prospects discourage participation in the labour market and therefore lower activity rate measures. This is illustrated, for example, in the reasons given by Patterson and Raffe (1995) to explain the higher incidence of staying on rates in Scotland. These are the expansion of higher education together with the progressive decline in labour market opportunities for 16-year old school leavers, reinforced by (regional) unemployment.

In order to more clearly illustrate the changing dynamics of the youth labour market, figure 2.3a summarises the main labour market destinations for the 16-years-old cohort in England and Wales between 1979 and 1992. The growth trend in the proportion of 16-year-olds remaining in full-time education is readily apparent, most noticeably in the 1990s. Also apparent is the emergence of youth training as a significant alternative school-to-work route in the 1980s peaking at more than 25 percent in 1987 before settling back and stabilising near current levels of around 10 percent (see also section 2.3.2). An even more dramatic, change is observable in the proportion of the cohort in employment, which fell dramatically in the early 1980s reflecting the onset of recession. In the early 1970s nearly two thirds of the cohort entered directly into employment, by 1992 this proportion had fallen to less than 10 percent.

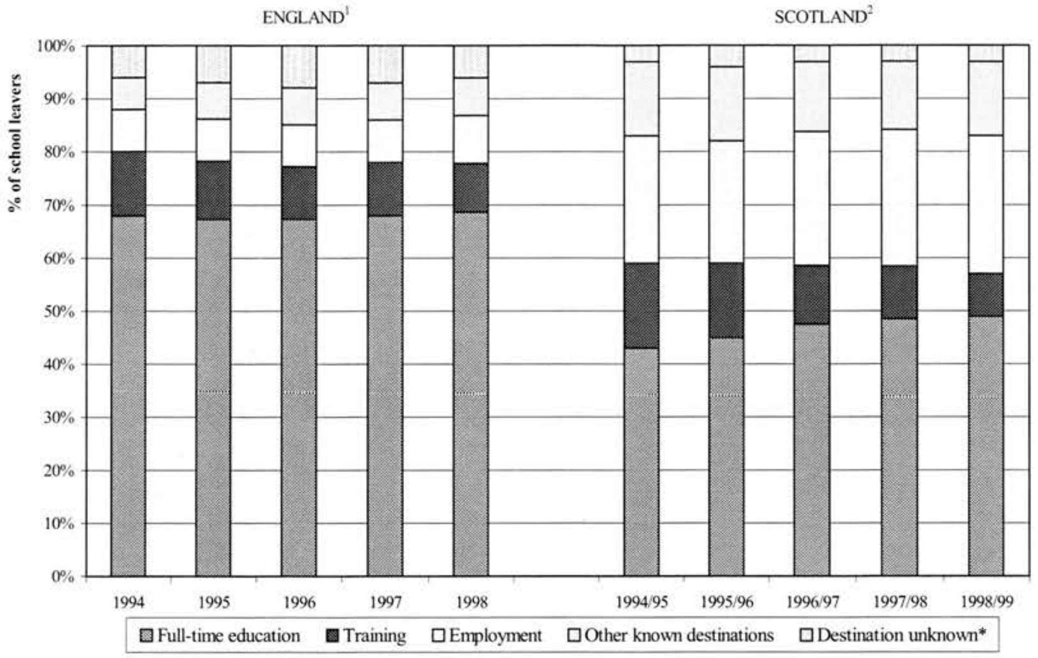
Figure 2.3a: Destinations of 16 year olds in England and Wales: 1979-92



Source: D*f*EE *News, various years*

Figure 2.3b shows more recent patterns in the destinations of school leavers in England and Scotland. These patterns are not directly comparable to those in figure 2.3a due to changes in the way data was compiled, although the English figures show a relatively stable pattern at levels similar to 1992 in figure 2.3a. The Scottish figures differ even more substantially in the way they are compiled with all senior secondary school leavers being included, whereas the English figures only cover year 11 school leavers. Nevertheless the trend to higher participation rates in full-time education, in this case amongst 16-18 year olds, is still readily apparent.

Figure2.3b: School leaver destinations in England and Scotland: 1994-1999



Source: *English figures – DfEE, Education and Training Statistics, (1997, 1998 and 1999 editions)*

Scottish figures – Scottish Executive, *Leaver Destinations from Scottish Secondary Schools (various editions, 1994-1999)*

Notes: *1. Figures for England relate to destinations of year 11 pupils (16 year-olds) leaving secondary school.*

2. These figures are not directly comparable to those for England as they cover the destinations of pupils from classes S4, S5 and S6 (16 to 18 year olds) who left school during or at the end of the year's academic session.

* *'Destination unknown' includes those who failed to inform the Careers Service or school what they were doing, and failed to respond to at least two attempts at follow-up by the Careers Office.*

2.2.3: Long-term Unemployment

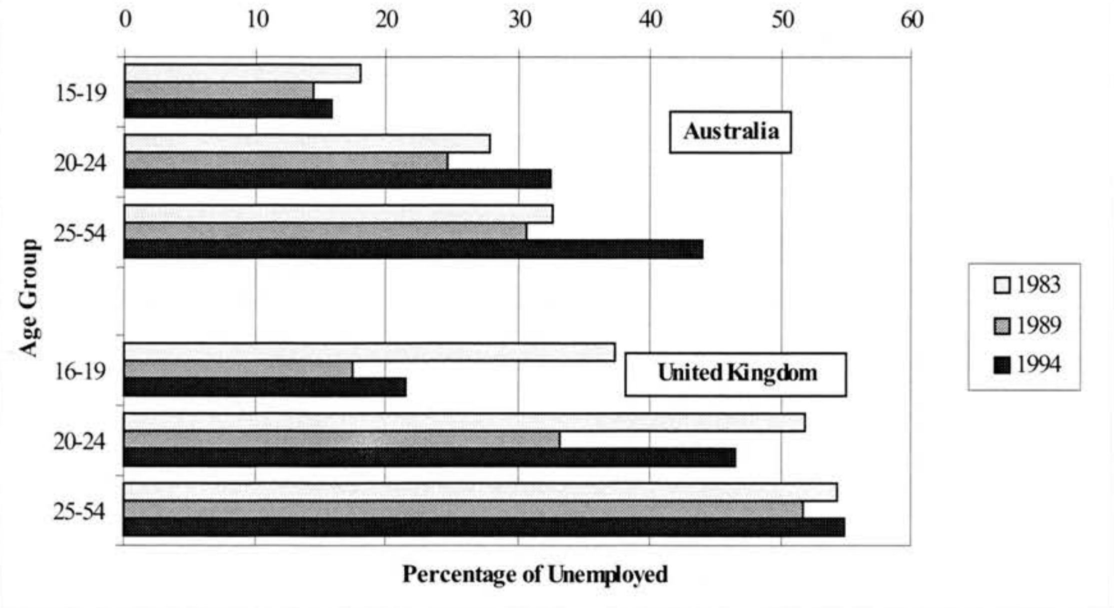
Another important aspect of labour market performance is the duration of unemployment, with many industrial countries witnessing increasing rates of long-term unemployment. Long-term unemployment is typically defined as a spell of one year or more. The key measure reported to represent the duration of unemployment is the 'incidence of long-term unemployment, which describes the likelihood (among unemployed people) of remaining unemployed for 12 months or more. Figure 2.4

reports the incidence of long-term unemployment in 1983, 1989 and 1994 for different male age groups in Australia and the United Kingdom. In general, the level for teenagers in both countries is relatively low and incidence increases with age. However, this needs to be kept in context with the different bandwidths for each age group and the trend in participation rates discussed above. More significant are changes in the incidence levels, which increased for all age groups with downturns in economic activity, as proxied by aggregate unemployment levels, over this period.

The significant fall in the incidence of long-term unemployment observed in the UK for 16 to 19 year olds post 1983 coincides with the introduction of the Youth Training Scheme (YTS – see the discussion on youth measures in section 2.3.2). A less pronounced fall is also evident in the reported unemployment rate of teenage males after 1983 and it continues on a modest decline for the remainder of the decade. The question as to whether this is enough to indicate the success of the scheme is addressed in Chapter 3 which describes the YTS in more detail and surveys reported estimates of its impact on the youth labour market. Both economies witnessed a slight improvement in the incidence of long-term unemployment between 1993/94 and 1999/2000, with the overall measures declining from 34.4 to 29.8 percent in Australia and from 31.9 to 30 percent in the UK (OECD, *Employment Outlook, 2001*).

In general, recorded long-term unemployment is higher for males, for example in Australia 67 percent of the long-term unemployed were males in 1989 declining to approximately 64 percent in 1999. This is due in part to women being more likely than men to leave the labour force when faced with unemployment. Moreover, the growth in the availability of part-time, casual and service industry employment in recent decades may have favoured the employment prospects of women (Austats, 2000). In general, no clear pattern is evident in the incidence of long-term unemployment. Although the incidence of long-term unemployment among young adults (20 to 24 years) appears to be subject to greater variation than the other age groups, particularly in the UK, which is consistent with the observed heightened sensitivity of youth rates to fluctuations in economic activity.[18]

Figure 2.4: Incidence of long-term unemployment amongst males in Australia and the UK by age for 1983, 1989 and 1994



Source: OECD Employment Outlook, 1996 July, Table 4.4.

Lindley (1996) sounds a cautionary note on the interpretation of duration statistics (and youth labour market statistics in general) observing that the:

> "...duration of unemployment has become so subject to recurrent movement between spells on schemes and temporary work that such statistics also fail to capture the extent of joblessness among those who have already left education and the difficulty of finding employment for those entering the labour market" (p. 167).

Given the context of this study any such blurring of statistical boundaries, through changes in policy need to be considered when assessing the relevant data. Moreover,

---

[18] For a detailed study of the likely causes and consequences of long-term unemployment see Payne, Casey, Payne and Connolly (1996). Their analysis uses data from the National Child Development Study (NCDS) to identify factors that increase the risk of long-term unemployment. Notably, for young men, these factors included poor qualifications and lack of formal job training.

the difficulty in determining the impact of interventions, such as training programmes, is increased when the intervention is part of a package of policies affecting alternative labour market options. For example, enhanced training programme provision is commonly accompanied by increasingly stringent restrictions to eligibility for unemployment benefits. This implicit compulsion clearly impacts on training participation patterns as well as the school-to-work transition process in general. However, the withdrawal of social security may also lead to a significant group of young people no longer appearing in government statistics after leaving the formal education system. This is an example of what has been labelled 'social exclusion'. The changing patterns in the school-to work-transition are discussed in more detail in the next section and with particular reference to Australia and the UK in chapter 3.

2.2.4: Demographic and Other Factors

A characteristic unique to the youth labour market is the considerable scale of new entry each year. All other things being equal this factor alone would result in higher levels of unemployment. The arrival of the baby boom generation in the labour market in the 1970s had significant consequences on youth employment opportunities and depressed wages for new entrants. Since the early 1980s, however, the size of the youth cohort has fallen relative to the 25 to 54 year old age group in most OECD countries. This trend is reflected in both Australia and the UK where the proportion of youth relative to adults fell by more than a quarter to a little over 30 percent over this period (see OECD, *Employment Outlook, 1996 July*, pp. 128-132).[19] In England the population of 16 year olds fell by more than 75,000 to approximately 600,700 over the decade between 1988 and 1998 (DfEE, 1999). As the preceding discussion has shown, however, the smaller youth cohort has not helped ameliorate the labour market problems facing young people. This suggests that any potential improvements were outweighed by other factors. Blanchflower

[19] In the UK it was widely predicted that the declining youth population would lead to significant distortions in employer recruitment patterns and drive up wages and thus reduce demand for education and training. Although labelled the 'demographic time-bomb' this predicted trend failed to materialise (Raffe et al., 1998).

and Freeman (1999) confirm that this pattern is repeated in most other OECD countries.

Korenman and Neumark (1997) considered the extent to which youth labour market problems may be ameliorated by demographic change. They examined the effect of changes in the relative size of the youth cohort and changes in adult unemployment rate (as a proxy for aggregate demand) on youth employment and unemployment in 15 OECD countries between the 1970s and early 1990s. They estimated that the response of youth employment to relative cohort size is close to zero. On the other hand, however they found that large youth cohorts lead to significant increases in the relative youth unemployment rate with estimated elasticity in the order of 0.5. That is a 10 percent increase in the size of the relative youth cohort will result in an approximate 5 percent increase in youth unemployment. An even more significant effect is found in the elasticity of the youth unemployment rate with respect to the adult unemployment rate estimated to be in the order of 0.8. Again this is consistent with heightened sensitivity previously observed for youth rates in response to fluctuations in economic activity. Korenman and Neumark also derive separate estimates for men and women and find that the relative cohort size effect for young men drops away almost completely while the adult unemployment effect increases to the order of 0.9. In contrast, the cohort size effect for young females is substantial with an estimated elasticity of 0.9 while the adult unemployment effect becomes less predominant with an estimate in the order of 0.6.

O'Higgins (1997) extends this form of analysis by differentiating between teenagers and adults. His results are broadly consistent with those of Korenman and Neumark again finding a greater sensitivity in youth rates to changes in the adult unemployment. However, with respect to the relative youth cohort size effect on youth unemployment, O'Higgins finds that it is predominately confined to the teenage sector with an estimated elasticity in the order of 0.52, compared with almost no effect for young adults. In summary, both the studies suggest that youth unemployment is much more responsive to general economic conditions than to changes in cohort size. Korenman and Neumark (1997) conclude that while demographic "...changes may yield some improvements in the youth labour market, more substantial reductions in youth unemployment will have to be generated from other sources" (p. 32).

Another significant dimension of the youth labour market is the increasing proportion of young people living in households where no one else is employed. The OECD (1998) reported that on average more than 20 percent of unemployed young people were in this situation and the proportion is growing. This phenomenon is particularly pronounced in the UK where the proportion of unemployed 15 to 19 year olds in households where no other person is employed grew from 26.6 to 32.4 percent between 1985 and 1996 (Blanchflower and Freeman, 1996). Other than the obvious welfare concerns this may affect the level of contact with the labour market and possibly job prospects. Payne (1987) estimated that in the UK, the likelihood of unemployment nearly doubled for youth living in households where the head of that household was unemployed.

A number of other factors are also important when assessing the state of the youth labour market. These range from the profile of school leavers in terms of educational qualifications, vocational training and any work experience they might have, to the relative wages for youth and the type of contracts on offer. The positive correlation between higher educational attainment and labour market success has been well established but this is not necessarily being achieved simply through higher staying-on rates. The increasing overlap of education and work has already been noted but this varies substantially between countries, as do wage systems and contractual arrangements. Further variation occurs across and within countries when ethnic, regional and other factors are included in the analysis. Such factors will hence be examined in more detail with relevance to our countries of interest, Scotland and Australia, in the following chapters.

The usefulness of employment and activity statistics is clearly limited to some degree because the nature and pattern of activity in the youth labour market are subject to considerable change (Lindley, 1996). Nevertheless, the trends common to most advanced industrial countries identified in this section indicate that the state of the youth labour market and hence the economic position of the average youth falls short of what is desirable with unemployment remaining high and persistent for many. In economic terms the cost of such unemployment is typically expressed in terms of forgone production and/or the loss of skills and income, but pervasive unemployment clearly has far wider implications.

For example Sen (1997, pp.160-165) identifies the following as different costs associated with unemployment, other than its association with low income: (1) loss of current output and fiscal burden (2) loss of freedom and social exclusion (3) skill loss and long-run damage (4) psychological harm (5) ill health and mortality (6) motivational loss and future work (7) loss of human relations and family life (8) racial and gender inequality (9) loss of social values and responsibility and (10) organisational inflexibility and technical conservatism. Although clearly interrelated Sen argues that the negative effects arising from these problems are "...cumulative, and they act individually and jointly to undermine and subvert personal and social life" (ibid. p. 160). Moreover, with respect to youth, a number of these problems are likely to be amplified and unemployment is likely to exact a 'particularly heavy toll' (ibid. p. 162). When considered together with the rising dependency ratio already being observed in many countries, due to ageing populations, the medium to long-term consequences of persistent youth unemployment are even more alarming.

Concern over factors such as these has prompted the OECD, the ILO and other organisations to recommend intervention in the labour market to help minimise unemployment and encourage skill development. The next section introduces the main categories of labour market intervention that exist in modern economies with particular emphasis on youth training.

2.3: Labour Market Programmes

The 1970s and 1980s witnessed a proliferation of labour market programmes largely in response to the adverse economic conditions associated with the 1970s oil shocks. The OECD focused attention on the importance of labour market policies by publishing data in *Employment Outlook* for public spending on, and participation in, labour market programmes from the late 1980s.[20] They use seven principal categories to classify policies that are focused on assisting the unemployed or those at risk of becoming unemployed. The first five categories, public employment services and administration, labour market training, youth measures, direct job creation and employment subsidies, and measures for the disabled are classified as *active* labour market policies. The remaining two categories, unemployment compensation and early retirement for labour market reasons, are then classified as *passive* labour market policies. This is by no means a complete classification of labour market programmes and the categories are also subject to some overlap.[21] The OECD (1990), responding to persistently high levels of unemployment in the 1980s, recommended that increased emphasis should be given to active labour market policy.

Government expenditure between 1985/86 and 1999/2000 on both active and passive labour market programmes are summarised for Australia and the UK in table 2.4, below. Youth and aggregate unemployment rates are also provided in the table in order to reflect the relative state of the economy. The figures are compiled annually by the OECD and expressed as a percentage of GDP. Table A2, in the appendix to this thesis, gives a breakdown of government expenditure on the same categories for 16 selected OECD countries while the average expenditures across these countries is reported in the final row of table 2.4.

---

[20] The definitions outlined here draw on the OECD's 1990 report, *Labour Market Policies for the 1990s* and chapter 2 of *Employment Outlook, 1993, July* where more detailed discussion can be found on all seven measures. The OECD reports have also spawned a raft of literature on the relative performance of ALMPs, see for example, Grubb (1994), Fay (1996), Janoski (1996) or O'Connell and McGinnity (1997). The OECD recently updated spending patterns between 1985 and 1998, by geographical grouping, see chapter 1 (pp.11-32) of *Employment Outlook, 2001 (July).*

[21] For example, a number of countries include measures for the disabled in the labour market training category and/or early retirement for labour market reasons in the unemployment compensation category.

Table 2.4: Public expenditure on labour market policy as a percentage of GDP in Australia and the United Kingdom, 1985 – 2000*

*Measure** / *Country*	Training			Job Subsidies	Disabled	**Total Active Measures**	U-benefit	Early Retirement	**Total Passive Measures**	Total Spending	**Youth U-rate**[a]	Aggregate U-rate
	PES & Admin	*Adult Measures*	***Youth Measures***									
	1	*2*	***3***	*4*	*5*	***(1-5)***	*6*	*7*	***(6-7)***		%	%
Australia [b]												
1985/86	0.11	0.02	**0.07**	0.19	0.00	**0.39**	1.31	..	**1.31**	1.70	**14.5**	8.0
1990/91	0.07	0.06	**0.04**	0.04	0.04	**0.25**	1.07	..	**1.07**	1.32	**13.3**	6.9
1995/96	0.24	0.15	**0.06**	0.31	0.07	**0.83**	1.28	..	**1.28**	2.10	**14.4**	8.5
1999/2000	0.20	0.02	**0.07**	0.11	0.05	**0.46**	1.05	..	**1.05**	1.51	**12.3**	6.6
UK [b, c]												
1985/86	0.14	0.09	**0.25**	0.22	0.03	**0.72**	2.01	0.05	**2.06**	2.79	**17.9**	11.2
1990/91	0.14	0.22	**0.18**	0.02	0.03	**0.59**	0.90	..	**0.90**	1.49	**10.1**	6.8
1995/96	0.20	0.10	**0.12**	0.02	0.03	**0.46**	1.33	..	**1.33**	1.33	**15.3**	8.7
1999/2000	0.13	0.05	**0.15**	0.01	0.02	**0.37**	0.58	..	**0.53**	0.94	**11.8**	6.1
OECD Avg. [b, d]												
1985/86	0.12	0.20	**0.11**	0.19	0.13	**0.75**	1.41	0.13	**1.54**	2.29	**17.2**	8.7
1990/91	0.13	0.24	**0.14**	0.14	0.15	**0.79**	1.25	0.09	**1.34**	2.14	**13.3**	7.1
1995/96	0.17	0.25	**0.13**	0.33	0.17	**1.04**	1.62	0.09	**1.70**	2.73	**18.1**	9.5
1999/2000	0.13	0.20	**0.11**	0.21	0.15	**0.82**	1.10	0.06	**1.16**	1.98	**14.1**	6.2

Source: OECD Employment Outlook, 1992-2001

[a] *Youth unemployment rate is for 16 to 24 years of age in Australia, 15 to 24 years of age in the UK.*

[b] *The figures for 1999/2000 are provisional only.*

[c] *All the measures for the United Kingdom exclude Northern Ireland.*

[d] *Unweighted average includes all 16 OECD member countries listed in Table A2, in the appendix to this thesis, excepting for the 1999/2000 average, which excludes Ireland and uses 1998/1999 figures for Canada, Japan and Portugal. Again the 1999/2000 figures are provisional.*

* *For notes pertaining to the different measures see footnote 22 in text.*

.. *Denotes nil or less than ½ of last digit used.*

No pattern is immediately clear from the tables with expenditure varying greatly between different countries and categories reflecting, among other factors, relevant economic conditions.[22] Closer inspection of the tables does however reveal that there is a shift toward expenditure on active labour market policy, which has generally increased in spite of economic conditions. This is reinforced by examining the portion of total expenditure assigned to active polices which has, on average, increased from approximately 32 percent in 1985/86 to approximately 41 percent in 1999/2000. In Australia, where total expenditure on labour market policies has tended to fluctuate markedly with changes in the aggregate unemployment rate, the share assigned to active polices has still increased from approximately 23 to 30 percent between 1985 and 2000. In the United Kingdom, total expenditure has declined sharply from 2.79 to 0.94 percent of GDP between 1985 and 2000 reflecting a general improvement in economic conditions, with aggregate unemployment falling from 11.2 to 6.1 percent. At the same time the share of total expenditure assigned to active policies has increased from approximately 25 to nearly 40 percent. The remainder of this section outlines the different active labour market categories in more detail.

2.3.1: Active Labour Market Policies

The public employment services and administration category generally includes a wide variety of placements, guidance and job search measures including related courses and counselling. It also includes various other measures of support such as aiding mobility and even light training measures. In the United Kingdom, for

---

[22] *Notes on measures in table 2.4:*

1. *Public employment services and administration* (PES & Admin) *includes: placement, counselling and vocational guidance; job-search courses and related counselling; support of geographic mobility to aid job search and all administration costs including benefit agencies and job centres.*
2. *Labour market training* (Adult Measures) *includes: training for unemployed workers and workers at risk for reasons of labour market policy.*
3. *Youth measures* (Youth Measures), *includes measures for unemployed and disadvantaged youth, support of apprenticeship and related forms of general youth training.*
4. *Direct job creation and employment subsidies* (Job Subsidies), *subsidies targeted to the unemployed. Grants to enterprises for capital costs are not included.*
5. *Measures for the disabled* (Disabled) *include vocational rehabilitation and work for the disabled.*
6. *Unemployment compensation* (U-benefit).
7. *Early retirement for labour market reasons* (Early Retirement).

example, it includes the interview based 'Restart' programme (the administration and evaluation of this programme is detailed in section 3.4). All administrative aspects of managing benefit agencies and job centres are also included in this category. This service should have the effect of reducing unemployment by facilitating job search and the filling of vacancies with increased efficiency through improved information and matching. The breadth of coverage in this category makes it difficult to determine any meaningful trend across the countries listed, in the first column of tables 2.4 and A2, other than the general increase in expenditure evident in the majority of countries featured in the tables.

Labour market training has the immediate impact of reducing unemployment by removing the participants from the labour market for the duration of the programme. If the programme is functioning well it should not only increase the probability of participants entering employment but also reduce any unemployment in their subsequent labour market experience. The individual should also benefit from higher future earnings and employers from enhanced productivity levels.[23] This would only be the case so long as the training adds to the skills and competencies of the participant. On a macro level increased training should contribute to the overall productivity and competitiveness of the economy. Youth training measures and apprenticeships should function in the same way and are discussed in more detail below. The expenditure on training measures reported in columns 2 and 3 of tables 2.4 (and the corresponding columns in table A2) is significant but shows a mixed trend. It increased in most countries between 1985 and 1990 but in a number, including the UK and Australia, expenditure has subsequently declined, particularly with respect to adult measures. Again this tends to reflect the state of the economy but may in part be due to the inconsistent and pessimistic results of evaluations carried out on many of the programmes in those countries. For example, in the UK, expenditure on youth measures peaked with large-scale youth training interventions in the 1980s (see discussion below) but expenditure has recently started to increase

---

[23] The standard economic approach views training as an investment in human capital where the individual gives up some proportion of income during the period of training in return for increased future earnings. For an accessible overview of issues relating to human capital investment, see Blundell et al., (1999).

again reflecting the government's commitment to the New Deal. In Australia expenditure on adult measures increased significantly, from 0.02 to 0.15 percent of GDP between 1985 and 1995 before subsequently declining back to 0.02 percent, while expenditure on youth measures has remained fairly constant at around 0.06 percent of GDP.

Direct job creation has largely taken place in the public sector and should have the impact of increasing employment directly. If the jobs were, however, only temporary then it would rely on the aspect of experience gained to increase the probability of participants subsequently gaining employment. Employment subsidies, on the other hand, are typically targeted at the private sector with the aim of reducing the cost of employment to employers. Again to be successful it relies on the experience gained by participants raising their productivity to the extent where it may lead to permanent employment. Expenditure on this measure, reported in the fourth column of tables 2.4 and A2, displays a very mixed trend typified in the contrast between Australia and the UK. In Australia, expenditure increased from 0.19 to 0.31 percent of GDP between 1985 and 1995 before falling to 0.11 percent in 2000. In contrast, expenditure in the UK fell from 0.22 to just 0.01 percent GDP over the same period indicating that emphasis was being placed on other measures in addition to improved economic conditions.

Measures to specifically aid the disabled include vocational rehabilitation and work experience. The levels of expenditure reported in column 5 of tables 2.4 and A2 tend to vary widely, however in most countries where expenditure is very low or zero the disabled are typically aided by measures in the other categories. On average there appears to be a trend of increased expenditure, independent of aggregate demand conditions in the economy, in order to make the labour market more accessible to disabled workers.

In summary, active labour market policies (ALMPs) can be defined as those that mobilise labour supply, strengthen job search and improve employment related skills. The objectives, targeting, implementation and funding of ALMPs vary widely across countries. In many countries, however, overall expenditure now exceeds more than one percent of GDP and the general trend is for this to increase. In contrast, many countries appear to be curtailing expenditure on passive labour market policies, unemployment compensation and early retirement for labour market reasons (columns 6 and 7 of tables 2.4and A2). Passive measures do, however, tend to be far more sensitive to fluctuations in aggregate demand and expenditure remains sizeable in most countries. Moreover, the success or otherwise of some ALMPs, such as training programmes, may be dependent on the different passive measures such as the level of unemployment compensation. If there is not a sufficient positive differential in remuneration between training and unemployment benefit, administrators may struggle to attract participants into their programme. Many countries, including Australia and the UK, have introduced an element of compulsion into ALMPs such as tying payment of unemployment transfers to participation in a suitable programme. Programmes that follow this approach are generally denoted as welfare-to-work programmes.

The marked shift in labour market policy from passive to active measures reflects an underlying belief that high unemployment and other labour market problems are neither temporary nor simply due to insufficient demand, but reflect underlying structural problems in advanced industrial economies (O'Connell and McGinnity, 1997). In Europe the emphasis on using active policies to address labour market issues has been reinforced by the European Commission's adoption of the *1998 Employment Guidelines*. In the case of the youth labour market the *Guidelines* state that:

> "Member States will ensure that every unemployed young person is offered a new start before reaching six months of unemployment, in the form of training, retraining, work practice, a job or other employability measure"

And in the case of the labour market in general they go on to state that:

> "Each Member State will endeavour to increase significantly the number of persons benefiting from active measures to improve their employability" [24]

---

[24] The full text of the *1998 Employment Guidelines* can be viewed on the Commission's website: http://europa.eu.int/com/employment-social/elm/summit/en/papers/guide2.htm

2.3.2: Youth Measures

Unemployed youth may benefit from a number of the active labour market programmes outlined above, particularly the public employment service, jobs subsidy and youth measures. The latter is of particular interest given the context of this thesis but also because of the investment element involved with potential long-term benefits to the economy. Youth training programmes are the most predominant of youth measures even though the OECD uses a relatively narrow definition of youth training focusing on measures to promote the transition from school-to-work, especially in groups that broadly correspond to upper secondary education. This definition excludes most cases in which comparatively young people take part in programmes available for any age group hence the expenditure in column 3 of tables 2.4 and A2 may be conservative. Moreover, a feature common in most publicly funded youth training programmes is that they are not exclusively limited to unemployed young people or other categories of disadvantaged youth but are open to any young person. In such cases the aim of the training programme may have broader aims such as improving the overall competitiveness of the workforce or to compensate for inadequacies in the existing education and training structure. The latter may include the provision of a 'second chance' for those who leave school prematurely through a lack of foresight or a high subjective rate of time preference (Friedlander, Greenberg and Robins, 1997).

In the United Kingdom the provision of youth training has been an important element in labour market policy since the 1970s and the introduction of the Youth Opportunities Programme (YOP). With the onset of recession in the early 1980s, considerable effort and resources were devoted to delaying the onset of, then ameliorating, the effects of a collapse of youth employment opportunities (Lindley, 1996, p.164). The provision of youth training has subsequently undergone considerable reshaping and now offers an increasing variety of choices in the school-to-work transition. For example, in little more than a decade the UK witnessed the transformation of the YOP, through three different permutations, into Youth Training. First, the Youth Training Scheme (YTS) launched as a one-year programme in 1983, then it was extended into a two-year programme in 1986, before being re-launched as Youth Training (YT) in 1989. YT brought with it an increased emphasis on National Vocational Qualifications (NVQs) with the accompanying 'Modern Apprenticeships', first available in 1994, increasing the emphasis on vocational qualification still further (see Deakin, 1996). This emphasis has been retained in the current schemes, National Traineeships (England and Wales) and Skillseekers (Scotland). The most recent initiative addressing youth issues is the Labour Government's New Deal for young people, a wide sweeping initiative which includes training as one of four possible options.[25] Significantly, these changes have all been introduced during an uncertain job market for school leavers and the withdrawal of social security benefits for most under 18s (Payne, 1995, p. 1).

Like the United Kingdom, Australia tends to fund youth training on a national basis although there are some differences in administration across different States and Territories. The main avenue of government support to training is through the subsidisation of wages for young trainees in employer-led (work-based) training, primarily apprenticeships and a complementary shorter programme of traineeships. These are buttressed by a number of smaller targeted programmes for 'at risk' groups, such as the Jobskills programme which operated between 1990 and 1994. The specific design of the Australian programme has evolved over time to conform to political, social and economic objectives of government including the promotion of training in non-traditional industries and occupations. The Jobskills programme and Traineeships, following substantive reform in the late 1980s and early 1990s, have a considerable number of elements in common with the UK's Youth Training (YT). This includes emphasis on qualifications, the linking of a vocational qualifications framework to training and some linking between welfare benefits and training for school leavers. Moreover, in 1988 so-called 'mutual obligation'

---

[25] The remaining options are either returning to full-time education, work in the voluntary sector or a place with the Environmental Taskforce, which is oriented toward environmentally friendly community projects. Twelve welfare-to-work pilot programmes were initiated throughout the UK in January 1998. The scheme allows for an initial 4-month 'gateway' for matching candidates to the most beneficial option, which may involve a 6-month subsidy or allowance being allocated, dependent upon the option chosen. The scheme was extended nation-wide in April of the same year and in January 1999it was further extended to include some unemployed persons aged 25 and over. The New Deal will be discussed further in chapter 3 of this thesis.

arrangements were introduced to help combat long-term unemployment. Under this initiative youth aged 18 to 24 years, after six months unemployment, are required to combine their job search activity with part-time work, voluntary work, education or training. These features were incorporated into the Modern Australian Apprenticeship and Traineeship System (MAATS), which assimilated the Jobskills Programme in 1994. MAATS had the dual objective of assisting the long-term unemployed and specific targeting of Aborigines and Torres Strait Islanders.[26] In 1998 the MAATS was relaunched in the form of *New Apprenticeships* that were designed to "...bring the flexibility of traineeships to apprenticeships and the status and recognition of apprenticeships to traineeships build on the flexibility of Traineeships combined with the integrity of apprenticeships" (Schofield, 1999b, p. 1).

In spite of such policy initiatives, however, the youth unemployment rate in Australia and the United Kingdom remains stubbornly high, as illustrated for the UK in figures 2a and 2b, above. Indeed there is some controversy as to whether active programmes, such as youth training, have any impact in combating unemployment or even improving the employment prospects of those who participate. Moreover, our understanding of the extent and coverage of training is unfortunately muddied by the different concepts and views of what constitutes training ranging from a formalised apprenticeship programme, to the informal learning by doing arrangements undertaken in many workplaces. Blundell, Dearden, Meghir and Sianesi (1999, p. 7) emphasise that studies using highly aggregated descriptions of training miss important differences in the determinants and effects of different forms of training. Significantly, existing evaluations of government assisted job training programmes have mostly treated the provision of training as homogenous. The results of these evaluations have led to widely varying inferences regarding the effectiveness of such programmes, particularly when non-experimental methods are employed.[27] This

---

[26] Australia also offers a plethora of smaller initiatives, which are discussed in chapter 4.

[27] For a summary of non-experimental evaluations of youth training in the UK see tables 4.1 and 4.3 in chapter 4.

variation in reported estimates of programme effect, to a great extent, underlies the motivation for this thesis.

In general active labour market policies differ widely in their objectives and, most probably, their impacts, both across countries and within countries over time. Ideally, they need to be examined in the context of the institutional structure of the labour market, including passive labour market policies and aggregate demand conditions. In the case of training evaluation education policies also need to be taken into account. In summary, ALMPs are only one element in a wide range of factors that impact on the employment prospects of young people. In a recent international survey of active labour market programmes by the OECD Labour Market and Training Section, youths were described as the most difficult labour market group to help. The study concludes that very careful targeting is needed and as a consequence relatively intensive and costly programmes may be needed to address problems in the youth labour market (Fay, 1996, pp. 28-29).

Programme evaluations attempt to determine the impact of various active programmes, both for the individual and on society at large. Individual impacts are usually measured in terms of employment probability and/or post-programme earnings. Social impacts include an estimation of any displacement effects associated with the intervention together with some accounting for externalities. The next section looks at the economic evaluation of such programmes in greater detail.

2.4: Programme Evaluation

In an age of constrained budgets and increasing accountability, evaluation of social programmes plays an integral part in programme monitoring and development. The OECD Secretariat defines evaluation as the "...systematic, critical examination of the objectives, implementation and impact of employment, training and social policies" (OECD, 1991, p.7). The generic goal of evaluation is however to provide feedback to a variety of interest groups including government (as sponsor), administrators, participants, voters and other relevant parties.[28] The OECD proposes that it is useful to subdivide this broad evaluation remit into three categories of evaluation activity. First, evaluation may focus on policy, its objectives and orientation or second, on the implementation of a programme, that is, assessing to what degree programme designs reflect the policy objectives. Third, evaluation may focus on the impact of a programme in terms of its policy objectives, which are not in themselves the object of assessment as they have already been established (OECD, 1991, pp. 7-8). It is the latter that economists are primarily concerned with as most studies seek to summarise the effect of the programme by examining post-programme outcomes and assessing whether the programme did, in fact, cause the outcome. Relative costs associated with the programme may also be estimated in order to apply a full cost-benefit analysis.[29] Following the classification of Trochim (1999), economic evaluation can thus be classified as a form of *summative* evaluation, which can adopt either microeconomic or macroeconomic methodology.[30]

The primary goal of macroeconomic evaluation studies is to examine the impact of the programme on aggregate employment and/or earnings (for example see Richardson, 1997). Therefore, they implicitly attempt to account for displacement effects, that is those situations where the effect of the labour market intervention is to reduce employment opportunities that would have otherwise existed, which is clearly

---

[28] A statement on the central importance held by evaluation for government policy in the UK can be found in *Labour Market Trends*, (November 1998, pp. 549-553) in an article which sets out the evaluation strategy, objectives and methods associated with monitoring of the New Deal. The Australian government's commitment to research and evaluation is set out in a document *Principles of Research and Evaluation* posted on the Department of Education, Employment, Training and Youth Affair's website: http://www.detya.gov.au. This document also articulates their ethical stance on evaluation design and practice.

[29] For an example of how a complete welfare analysis might be carried out see Heckman and Smith (1998), who investigate the impact of imposing different social welfare functions on outcomes in policy evaluation. In practice most applications in the evaluation literature report only the mean programme impact, which is estimated ignoring costs, transfers and distributional issues, see chapters 3 and 4 in this thesis.

[30] Trochim (1999) distinguishes between *summative* and *formative* evaluation. The latter are defined as evaluations that strengthen or improve the programme being evaluated by examining aspects such as implementation, organisation and procedures. The first two evaluation activities defined by the OECD are thus formative.

a difficult task given the many extraneous factors that affect the outcome.[31] On the other hand, at the micro level evaluation studies examine the impact on individual participants, which is complicated by the absence of valid comparison groups. In other words, both approaches rest on the quality of the empirical input used to generate their estimates and on the widespread acceptance of causal identifying assumptions (Heckman and Smith, 1995, pp. 85-86). Moreover, there may be further externalities arising from the programme that neither micro nor macroeconomic studies pick up directly, such as reductions in crime and healthcare costs associated with the long-term unemployed (Fay, 1996, p. 7).[32] This thesis is primarily concerned with the microeconomic evaluation of youth training programmes. The advantage of this approach is that it can give a specific idea of what type of individuals benefit from the relevant training programme and by how much. Moreover, this more fundamental query needs to be established before we endeavour to employ more elaborate structural models that place increased demands on data collection and availability.

The assessment is quantitative and causal in nature and involves the construction of a counter-factual condition against which programme performances can be measured. That is, we need to first establish what the outcome would have been for a programme participant if they had not participated in the programme. The difference between the observed outcome and the counterfactual outcome is then a measure of the impact of the programme for that participant. The counterfactual can be established using either non-experimental or experimental methods. Only a brief overview of the two approaches is provided here as the relative merits and limitations of each approach are examined in more detail in the next chapter.

---

[31] The displacement effect is made up of two components: the substitution effect and a dead-weight loss. For example if the intervention was one of subsidised youth employment, then the substitution effect would relate to those situations whereby the effect of the intervention is to reduce employment in another group, for example female part-time employees, because the subsidised youth placements are used instead. The dead-weight loss would relate to the situation where existing youth jobs are simply converted to the subsidised placements.

[32] See Freeman (1996) for an investigation of the relationship between crime and unemployment among young males in the USA. He has carried out extensive research in this area and his studies have found evidence that any deterioration in economic opportunity is likely to be accompanied by an increased involvement in crime.

Non-experimental methods typically employ strategies that compare participants and non-participants based on their observed characteristics such as age, gender, qualifications, location, etc. Econometric techniques are employed to estimate the counterfactual for participants based on the outcomes experienced by non-participants with the same or substantially similar observable characteristics. It is virtually impossible, however, to establish satisfactory matching on characteristics such as motivation or attitude, which among other things are extremely difficult to measure. Therefore the evaluator can never be completely confident that any observed post-programme difference between participants and non-participants is attributable to the programme or pre-existing unmeasured differences between the two groups. If the latter is the case then the study will suffer from what is known as *sample selection bias*, which can lead to biased inferences about the impact of the programme.

Selection bias is at the root of the evaluation problem, which arises because one cannot simultaneously observe the same person in a programme and out of it. Hence, it is not possible to determine the programme impact for any individual person. With appropriate micro data, however, it is possible to construct the outcome distributions for participants, non-participants and the population as a whole. It is also hypothetically possible to construct the outcome distributions for the relevant counterfactual. If this is possible and the outcome distribution for non-participants is found to coincide with the counterfactual for participants, or can be adjusted to do so, then the difference in the means of the participant and non-participant distributions is the mean impact of the programme (Heckman, Smith and Clements, 1997). A formal statement of the evaluation problem is presented in chapter 3, which also examines how counterfactuals might be established.

Controversy has arisen over the reliability of non-experimental estimators because reported estimates on the mean impact of similar training programmes have tended to be widely divergent. Research in the USA has shown that this is largely dependent on the methodology employed. Given that different evaluation procedures make different assumptions about the distribution of differences between participants and non-participants it is not altogether surprising that they produce different estimates of programme impact. Hence, it is important to assess the degree of variation due to different non-experimental approaches in different settings. For policy requirements, however, it is often sufficient to simply identify the direction of impact rather than its precise size. Therefore, in general, if estimates derived from non-experimental methods are in agreement and appear to be accurate indicators of the direction of impact; they may be used with a reasonable degree of confidence. In cases where they fail to agree even on direction, it may be possible to judge which estimates are more reliable through the application of econometric testing, however, one of the controversies that continues to hinder non-experimental methods is the absence of any reliable means for choosing the appropriate estimator. However, the application of sensitivity analysis to non-experimental studies may help identify the conditions under which different non-experimental estimators can be successfully employed and is hence utilised in chapter 6 of this thesis.

Experimental methods, on the other hand, select individuals from the population of interest and randomly assign them to either a *treatment* group that takes part in the programme, or to a *control* group that is not permitted to participate. Ideally this approach eliminates any systematic differences between participants and non-participants, allowing an accurate counterfactual to be established for assessing the programme impact. Experiments are however costly, politically unpopular on ethical grounds, and difficult to implement in a way that produces valid impact estimates of ongoing programmes. They may also be subject to bias through the practical difficulty of maintaining a valid control group. A potentially more serious problem is *randomisation bias* that occurs when the type of person that participates in the experimental component of the programme is different to the type of person who would participate in the programme, as it would normally operate (Heckman and Smith, 1995). These and other related factors are discussed in greater detail in chapter 3.

2.5: Chapter Summary

This chapter has provided the necessary context for the research conducted in this thesis by assessing the state of the youth labour market in our countries of interest, Australia and the United Kingdom. This is achieved by developing a statistical profile of the respective youth labour markets focusing on the unemployment rate, the labour force participation rate, and the employment/population ratio. It is shown that, in general, the state of the youth labour market falls short of what is desirable with youth unemployment remaining high and persistent in both countries. Alarmingly, youth unemployment rates remain high relative to total unemployment despite the fact that significant proportions of teenagers are choosing to stay on at school and more are leaving with qualifications. Moreover, the pattern is heightened for young men and, in the case of Australia, for indigenous youth. The same groups are more likely to experience long-term unemployment. The trend of higher relative unemployment among young people is observed in a wide range of advanced industrial countries at the same time as their labour markets are subject to ageing populations and rapidly changing skill demands.[33] Together these factors have ensured that the youth labour market is placed high on the policy agenda and that intervention is seen as a realistic, if not necessary, option.

In the face of increasingly adverse economic conditions in the 1970s most industrial countries responded by increasing levels of intervention in the labour market. This chapter has examined the pattern of intervention, in terms of expenditure on labour market programmes, in a selection of OECD countries between 1985/86 and 1999/2000. Although subject to economic fluctuations, a marked increase in the proportion of expenditure on 'active labour market policies' is generally observed in member countries, including Australia and the UK. Attention was focused on *youth measures*, predominately youth training, and it was illustrated that this is an active policy area that may have broader aims than simply absorbing fluctuations in youth unemployment. For example, intervention may be motivated by long-term goals, such as improving the overall competitiveness of the workforce or to compensate for inadequacies in the existing education and training structure. This was illustrated by describing the youth measures undertaken in the UK and Australia, which have both sought to increase levels of youth training.

---

[33] As detailed in the text above, additional youth labour market statistics are provided for a wider selection of industrial countries in the statistical appendix to this thesis (table A1).

This chapter has also introduced the key approaches adopted in economics to the evaluation of interventions, such as youth training programmes, and clarified a number of key definitions used in the analysis throughout this thesis. The next chapter begins by investigating evaluation issues in more detail before critically examining the relative merits of the two main microeconomic approaches to social programme evaluation, that is, experimental and non-experimental methods. The chapter then concentrates on the latter and examines the main theoretical advances that have been made in this field including an examination of various methods that might be used for choosing between different non-experimental estimators.

Chapter 3: Evaluation Issues

3.1: Introduction

The economic evaluation of employment and training programmes is central to manpower policy decision-making; nevertheless, considerable controversy accompanies evaluation research and reporting. Official agencies in most countries, including the United Kingdom and Australia, typically continue to report simple placement rates, that is, the percentage of participants in a programme who obtain jobs on the completion of the programme, or changes in the rate of unemployment, as measures of success of their programmes. In the United Kingdom for example, the Department for Education and Employment (D*f*EE) announced in November 2000 that the New Deal had placed 250,000 young people in subsidised and unsubsidised work (D*f*EE Press Notice, 2000/0556).[34] Measures of 'gross effect', such as these, are generally useful as a rough guide for the monitoring of a programme's performance and may even assist the examination of a particular programme's relative cost efficiency. They cannot, however, isolate what impact is actually attributable to the programme and hence, are no substitute for detailed economic evaluation. Without reliable impact measures for past and current policies reliable assessments of future policy options cannot be made. Microeconomic evaluation attempts to assess the direct impact of a programme on participants (or groups of participants) therefore, as outlined in chapter 2, an effective evaluation needs to be able to establish a feasible *counterfactual* to that programme.

The counterfactual asks, "what would have happened to the participants (or labour market, economy, etc) in the absence of the programme, all things being equal?" (Ryan and Buchtemann, 1996, p. 312). This introduces a number of unique challenges for the analyst. The absence of an observable counterfactual outcome for

---

[34] This total includes placement rates in England and Wales only. Figure 4.2 in chapter 4 shows the reported destinations of young people after leaving government assisted youth training in the UK in terms of being in employment, unemployment or a 'positive outcome'. The Department for Education and Employment (D*f*EE) defines a positive outcome as being in work, education or entering a further training programme. The Department for Education, Employment, Training and Youth Affairs (DETYA) employs a similar definition in Australia.

participants in a training programme motivates the use of outcomes for non-participants to infer what the outcomes would have been for participants had they not taken part. The difference between the regression-adjusted outcomes in these two states is then often assumed to be the impact of the programme. This type of impact estimate, common in early evaluation literature, is however likely to be biased due to pre-existing differences between participants and non-participants. These pre-existing differences are due mainly to choices made by participants; programme administrators and/or the analyst, giving rise to the primary evaluation problem or *sample selection.*

Programme evaluation in the youth case is particularly difficult because most of the sample population do not have labour market experience in the period prior to programme entry. This increases the complexity of identifying pre-existing differences while growth in maturity may impact on differences in outcomes even in the absence of the programme. Evaluation is further complicated as policy objectives may vary with training programmes making it crucial that any evaluation be carried out against the stated objectives of the specific programme. However, objectives are often not clearly defined and programmes may also lack specific focus, incorporate multiple objectives and/or overlap with the aims of other interventions. Moreover, within the lifetime of a programme, the objectives may appear to change, different interest groups, or *stakeholders*, may disagree over what the actual objectives are and funding provisions may change. Consequently, without specific guides researchers must assess what would seem like a reasonable indicator in the light of broad programme aims. Evaluation tools thus not only need to be able to establish a credible counterfactual but also address a variety of questions of interest.

In Section 3.2 the concept of sample selection is outlined in intuitive terms before proceeding to a formal statement of the evaluation problem. This section shows that measures of gross impacts, such as placement rates, and standard regression analysis are not adequate measures of programme effectiveness mainly due to sample selection. It is shown that failure to take sample selection into account may introduce a significant bias into the results of evaluation studies. Factors that are likely to give rise to the evaluation problem in the training context are discussed together with the extent of the problem in economics. The remainder of this chapter then examines the methods developed by economic analysts in order to solve the evaluation problem while the main empirical findings of recent evaluations of government sponsored youth training programmes in the UK and Australia are reported in a survey of the relevant literature in chapter 4.

A wide range of methodologies has been employed to assess the impact of training programmes (and other active labour market policies), however we can identify two distinct approaches to solving the evaluation problem. The first of these relies on data from social experiments. Section 3.3 looks at how the counterfactual is established in this case. Experimental methods use random assignment in order to create a comparison group that is not systematically different from the participants. Hence experimental impact assessments, aside from sampling error, should reflect only the effects of the training programme being evaluated (Orr, 1999, p. 13). Some analysts view social experiments as the only valid evaluation method and hence, in section 3.4, some limitations to the practical application of social experiments are discussed.

In the second approach, presented in section 3.5, the comparison group is chosen by the evaluator using non-experimental data and econometric techniques in order to construct the counterfactual. In the absence of random assignment to the programme, the decision rules governing participation need to be modelled as well as the relevant outcome functions. The decision to participate is not however generally observed by the evaluator. Hence, assumptions are required to identify the model and because empirical impact assessments are sensitive to these assumptions, there is inevitably scope for disagreement in the causal interpretation of micro data in the social sciences (Heckman and Smith, 1996, p. 52). Section 3.5.1 looks at the choice of comparison groups, followed by a formal presentation the conventional Heckman 'two-step' selectivity model and some commonly applied variations, while section 3.5.3 examines extensions to the basic model.

Together, sections 3.4 and 3.5 establish a case for continued research into the application and behaviour of non-experimental methods based partially on the assertion that, in practice, experimental methods are not the panacea that some analysts hoped for and partially on advances that have been made in non-experimental methodology. Even in experimental settings non-experimental methods are still required to address many important questions about the impact of training (Heckman, 1992; Ham and LaLonde, 1994). Moreover, given the relative paucity of experimental data and restrictions on the application of social experiments in many countries, including the United Kingdom and Australia, it is crucial to validate applicable non-experimental methods and accurately interpret the estimates derived.

Due to the wide range of applications requiring efficient evaluation within the social sciences and growing recognition of evaluation bias across the discipline we have witnessed a flux of alternate non-experimental methods. The flexibility of these different methods allows different criteria of interest for estimating programme impact to be assessed and these are explored in section 3.5.4. Nevertheless, while it is clearly illustrated that major advances have been made in non-experimental evaluation methods, it is not clear that we have the necessary tools to obtain unbiased impact estimates of training programmes or other labour market interventions. In fact one of the major criticisms levelled at the non-experimental method is that it is difficult to choose among the alternate non-experimental estimators (Orr, 1999, p. 13; LaLonde, 1986, p. 614). This problem is examined in Section 3.6 where three different approaches for evaluating the validity of different non-experimental methods are discussed (sub-sections 3.6.1 to 3.6.3).

First, the use of data from social experiments to evaluate the likely reliability of non-experimental methods by comparison of outcome assessments is considered in section 3.6.1. Following the formative work of LaLonde (1986) the comparative approach has received considerable attention in the literature and has generally concluded that non-experimental methods are inadequate. This approach rests on the crucial assumption that the experimental evaluation is unbiased. In practice, however, social experimentation is itself subject to a number of limitations as identified in section 3.4 and, as such, it may not be an appropriate benchmark for assessment of non-experimental methods. Moreover, this approach is limited to those countries that allow social experiments to be conducted.

Then, in section 3.6.2, we consider the use of simulated data to assess the validity of non-experimental methods in different settings. This approach has considerable flexibility making the testing of a wide range of different assumptions and parameters plausible, however, results have been mixed and the literature generally adopts a neutral stance suggesting caution is exercised in the application of non-experimental methods.

The third approach we consider, in section 3.6.3, is sensitivity analysis, where different model specifications are applied to a common database to reveal more about the construction and interpretation of the resulting estimates as well as identifying any bias introduced by alternate model selection. This approach has been usefully applied in a number of labour market settings, most notably by Mroz (1987) who used it to assess the sensitivity of an empirical model of married women's hours of work to different economic and statistical assumptions. This is the approach adopted in the empirical section of this thesis (see chapter 6 below).

Section 3.6 also reviews some of the more significant applications of each approach in the relevant literature. The evaluation of alternate non-experimental methods may aid choice of the appropriate model for the relevant programme and the available dataset. It should be stressed however, that the aim is not to identify a single all-encompassing model; Heckman and Smith (1996, p. 54) argue that the relevant literature has been marred by such claims. Rather the objective is to enhance our understanding of the operation and applicability of non-experimental methods in different settings. This includes examination of the trade-offs made, through the adoption of different identifying assumptions, and the justification of the different assumptions necessary for evaluation of the impact of training on outcomes given the likely presence of sample selection bias. A chapter summary is provided in section 3.7.

3.2: Evaluation Bias

The primary source of evaluation bias in the assessment of training programmes is sample selection. Sample selection bias encompasses any differences between the programme participants and the non-participants that affect the outcome of interest. It occurs when a rule other than simple random sampling determines the sample gathered from the underlying population of interest. For example, in the case of training, observed outcomes are not randomly assigned across the population but are a result of individual choices, circumstances and attributes. If we ignore this endogeneity our estimates of the returns to training may be biased. In other words, the sample is *truncated* and cannot be relied upon to accurately represent the true population distribution of characteristics, regardless of sample size.[35]

Although sample selection is a ubiquitous problem in the social sciences and arises in many different settings, ranging from education to immigration, formal analysis in economics is a relatively recent development. It received considerable attention in the econometric literature of the 1970s resulting in the development of several methods to model the selectivity and correct for the bias, see for example Gronau (1974), Lewis (1974) and Heckman (1974).[36] These early studies tended to focus on determinants of wage and labour supply behaviour of women but the methods developed have subsequently been applied to a wide range of issues including, the impact of unions, housing demand, returns to education, migration and the impact of various social programmes.[37] With respect to training programmes Moffitt (1987, p.

---

[35] Formally, sample selection is a form of censoring where the point of truncation of the dependent variable is a function of another variable (the selectivity variable) and hence it is also sometimes described as *incidental truncation* (Goldberger, 1981; Greene, 1995). Breen (1996, pp. 2-4) discusses the confusion that sometimes arises in the literature between censoring and truncation definitions. See also section 3.5.

[36] James J. Heckman was awarded the 2000 Nobel Laureate in Economics for his development of theory and methods for analysing selective samples. He shared the award with Daniel L. McFadden who was honoured for his development of theory and methods for analysing discrete choice.

[37] In the UK the most intensively evaluated social programme is the Youth Training Scheme, the results are surveyed in Chapter 4. Other UK studies include King (1980) who examined the issue of housing demand (to rent or purchase); Murphy, Sloan and Blackaby (1992) who analysed the relationship between union membership and earnings; and Dolton and Makepeace (1992) who looked at the returns to education (graduate earnings). Most of these apply some form of

154), in his overview of evaluations in the USA, cites Ashenfelter (1978) and Bloch (1979) as the first analysts to directly address the problem of sample selection bias. However, widespread suspicion over the reliability of econometric methods in social programme evaluation, at times exacerbated by contradictory and/or counter-intuitive findings, encouraged the adoption of social experiments in some countries (discussed below in section 3.3).

There are several potential sources of sample selection bias in databases pertaining to training programmes. First, individuals making choices to belong to one group or another typically generate the data used for evaluation that is they are *self-selected.* Roy (1951) provided an early discussion of the problem of self-selection positing that individuals choose between discrete 'occupations' (hunting and fishing) dependent on their comparative advantage in each. That is, self-interest drives them to choose the occupation that produces the highest utility (income) based on their own latent (unobserved) skill base. The observable skill distributions are thus the outcome of the unobserved 'selection rule' adopted by the individuals.[38]

Second, decisions made by analysts also impact on the sample, for example it is common to analyse only those observations followed for the full length of the sample, that is, to ignore any attrition in the analysis. Attrition is an inherent problem in social science data, whether generated from social experiments or non-experimental sources, and if the attritors systematically differ from those remaining in the programme estimates will again be biased. Similarly, it is also possible programme administrators may give preference to more able applicants to enhance the performance of the programme. Although some of the administrators' judgements may be based on variables that are observable to the evaluator, some may be based on unobservable factors such as the administrators' assessment of the individual's motivation, reliability, and potential. Such procedures may be viewed as having the same effect on structural estimates as self-selection (Heckman, 1979, p.

Heckman's (1976 and 1979) two-step method, discussed below in section 3.5, which corrects for any bias by treating it as an omitted variable.

[38] For detailed discussion of the empirical content of the Roy model see Heckman and Honore (1986) or Maddala (1983, pp. 257-258).

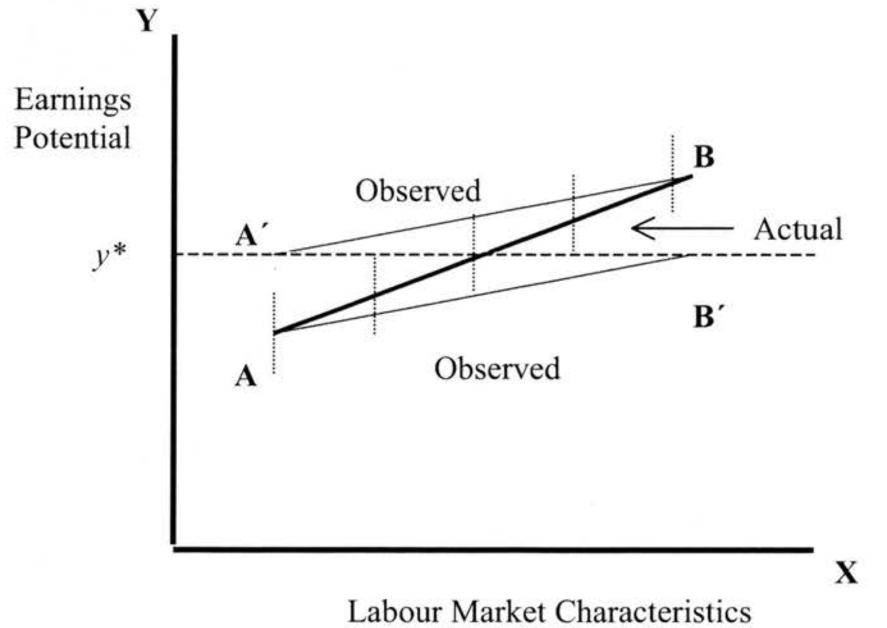
154). Even in the case of social experiments individuals may exercise different motives for joining or leaving a programme and the administrator may exercise judgements that result in a bias.

The essence of the evaluation problem is that those who participate and remain in a programme may be different in some important but unmeasured way to those who do not participate in the programme. For example, individuals entering a programme may be more motivated, and therefore have greater employment and earning potential, than those who choose not to enter the programme. In such a case the returns to training will be overstated. Equally individuals who are less optimistic about their labour market prospects may enrol in job training programmes in disproportionate numbers. If their pessimism is based on a realistic, but unobserved, assessment of their opportunities, we should expect their outcomes in the absence of training would be lower than those of individuals with identical observable characteristics who do not enrol in training.

Sample selection can be graphically illustrated for the simple case, as in figure 3.1 adapted from Main and Shelley (1990, figure 1), where the bold line **AB** reflects the expected outcome *Y*, say earnings potential, of randomly chosen school leavers as a function of some labour market characteristic *X*, such as age gender and/or the local unemployment rate. The vertical broken lines represent variation about the expected outcome owing to unmeasured characteristics, such as motivation (assumed to follow a uniform distribution).

Assume that more motivated individuals can be distinguished from less motivated, or more pessimistic, individuals by assigning a potential post programme output measure greater than y^* to the former and a measure less than y^* to the latter. The line **AB** hence reflects the output potential of individuals in the absence of the programme, that is, the true counterfactual of the programme. If however, some mechanism filtered entrants to the programme so as only the more motivated participated then the estimated outcome measure for the remaining non-participants would be the line **AB′**, a biased estimate of the benchmark state **AB**. Conversely, if only the less motivated individuals participated in the programme then the estimated outcome measure for the remaining non-participants would be the line **A′B**, a bias in the opposite direction.

Figure 3.1: Sample Selection Bias*



The uncertainty about the presence, direction, and potential size of selection bias is often used as an argument in favour of social experiments (for example, see Burtless, 1995, p. 72; Maynard, 1993, p.102). Meyer (1996) however, lists sample selection as a continuing threat to both internal and external validity of experimental estimates. While Heckman and Robb (1986) choose to view it as simply a special case of the general problem of drawing inferences from observational, non-experimental data of the kind typically available in social science research.

In terms of applied evaluation the bias arises when the unmeasured influences on the outcome, typically measured as earnings or employment status, are correlated with the probability of participation, that is non-random assignment. Hence, in the presence of sample selection, a comparison applying the conventional estimation technique, ordinary least squares (OLS) to the outcome equations of participants and non-participants may result in inconsistent and biased estimated coefficients rendering suspect any conclusions drawn from the resulting outcome estimates. In order to remove the bias we need to compare the outcome for an individual participant in a training programme with the outcome the same individual would have received had they not participated.[39] In other words compare the measured outcome with the counterfactual outcome for the same individual.

To more clearly illustrate how the evaluation problem impacts on both experimental and non-experimental methods we assume, following Heckman, Smith and Clements (1997), that there are two potential states for all individuals, y_1 and y_2. Where y_1 is the outcome obtained given participation in the training programme being evaluated and y_2 is the hypothetical outcome, or counterfactual, in the benchmark state of non-participation. These outcomes are conventionally earnings or employment in the two states and the effect of the programme is defined as a comparison of y_1 and y_2, such as $(y_1 - y_2)$. If it was possible to observe y_1 and y_2 for all individuals the evaluation problem would not exist and the mean impact of the programme $(y_1 - y_2)$ could be found for each individual and various populations of interest.[40]

The evaluation problem arises precisely because we do not observe y_1 and y_2 for all individuals and, thus, it is not possible to estimate the joint distribution of (y_1, y_2) or the distribution of gains directly. Instead, a statistical approach is adopted that focuses on estimating the joint distribution across the population conditional on the

---

[39] Heckman and Robb (1985, p. 161) note that this comparison actually leads to two distinct questions. Assuming the outcome of interest is earnings the first asks, "What would be the impact of training on earnings if people were randomly assigned to training?" which standardises the analysis to a random individual (with given measurable characteristics) drawn from the population sample of all participants and non-participants. This is the appropriate question if the focus of attention is on the average (population) impact. The second asks, "How do the post-program earnings of the trained compare to what they would have been in the absence of training?" which standardises the analysis to the population of participants only. That is, the effect of "treatment on the treated". This is the appropriate question if interest centres on forecasting the average impact of training on trainees when the same selection rule(s) are retained for future trainees. These questions will only have the same answer when training has the same impact on everyone or there is random assignment to training.

[40] The mean impact is the conventional measure of interest in both experimental and non-experimental studies, with particular focus on $E(y_1 - y_2 | d = 1)$ where d denotes participation, that is the estimated effect of 'treatment on the treated'. Burtless (1995, p. 70) and Heckman (1992, p. 216) argue that this is largely due to the fact that people untrained in statistics, such as politicians and programme administrators, more readily understand means. Further criteria of interest are explored in section 3.5.4.

available information or some features of it. For the purposes of clarity and notational convenience however, at this time, we ignore the dependence of the following measures on the relevant explanatory (X) variables and the entire analysis can be assumed conditional upon them.

Allow d to denote participation in the training programme, where $d = 1$ if an individual participates and 0 otherwise. Then it is possible, using microeconomic data of the form typically available in the social sciences, to determine the conditional outcome distributions across individuals:[41]

$$F(y_1 \mid d=1) \qquad \text{(for participants)} \qquad (3.1a)$$
$$F(y_2 \mid d=0) \qquad \text{(for nonparticipants)} \qquad (3.1b)$$

In the case of universal coverage of the programme we would not observe the non-participant state (3.1b) because $d = 1$ for everyone. Even when programme coverage is partial we cannot observe y_2 for participants or y_1 for non-participants. Hence, without additional information it is not possible to construct the counter-factual conditional distributions. Either, what participant outcomes would have been had they not participated, that is:

$$F(y_2 \mid d=1) \qquad (3.1c)$$

or what non-participant outcomes would have been had they participated, that is:

$$F(y_1 \mid d=0) \qquad (3.1d)$$

Neither is it possible, due to the inability to observe (y_1, y_2) for either participants or non-participants, to construct the conditional joint outcome distributions:

$$F(y_1, y_2 \mid d=1) \qquad (3.1e)$$

and $$F(y_1, y_2 \mid d=0) \qquad (3.1f)$$

Unless participation in the programme is random with respect to outcomes, so that $F(y_2 \mid d=1) = F(y_2 \mid d=0)$, it is not possible to use the conditional outcome

---

[41] For ease of interpretation we assume F is a normal distribution but the analysis would equally apply to any symmetric distribution.

distributions (3.1a and 3.1b), derived from non-experimental data, to estimate conventional parameters of interest.

Using the population means for participants and non-participants we can, however, obtain the following expectation:

$$E(y_1 \mid d=1) - E(y_2 \mid d=0) = E(y_1 - y_2 \mid d=1) + [E(y_2 \mid d=1) - E(y_2 \mid d=0)] \quad (3.2)$$

Where the final term on the right hand side (in square brackets) captures any selection bias on the basis of the non-participation state y_2. As non-participants typically differ from participants in the non-participation state this is generally not zero.

If experimental data was available sample selection can be resolved as social experiments can produce $F(y_2 \mid d=1)$ under certain ideal conditions, discussed below. Social experiments cannot, however, recover the non-participant counterfactual $F(y_1 \mid d=0)$ as one cannot force non-participants to participate and the conditional joint outcome distributions cannot be recovered without invoking additional assumptions. The latter is due to the fact that we cannot observe (y_1, y_2) for any individual and it is only possible to determine features of the joint distribution that depend solely on $F(y_1 \mid d=1)$ and $F(y_2 \mid d=1)$, even when experimental data is available. In summary, all evaluation methods require assumptions to replace missing data at some point in their estimation procedure. Heckman (1992, p. 213) argues that the assumptions in the experimental case differ from but are not necessarily any more plausible than those made in the non-experimental case. The next section looks at the experimental method in more detail.

3.3: Experimental Methods

The case for conducting social experiments, or more accurately 'field experiments', to evaluate active labour market programmes and training programmes is generally based on the perception that they only require a small number of qualifications (Burtless, 1995; Barnow, 1987; LaLonde, 1986). The successful evaluation of training programmes using social experiments rests on the random assignment of eligible participants to treatment (individuals who participate in the training) and control groups (individuals randomly denied the training). It is then assumed that in sufficiently large samples any underlying differences between participants and controls are trivial and hence, the unadjusted difference in mean outcomes between the two groups is defined to be an unbiased estimate of the programme impact. Randomised social experiments with appropriately chosen control groups are thus specifically designed to eliminate the correlation of any unobservable differences across individuals with the returns to training. In other words, remove the problem of sample selection (see Burtless, 1995, pp. 68-72).

To illustrate how experimental methods are designed to overcome the evaluation problem we retain the assumption that there are two potential states (y_1, y_2) for all individuals. Where y_1 is the outcome obtained given participation in the training programme being evaluated and y_2 is the hypothetical outcome, or counterfactual, in the benchmark state of non-participation. Following Heckman and Smith (1996) two new variables are introduced to clarify the necessary process for construction of an experimental control group comprising individuals who would participate but are randomly denied access to the programme. Allow d^* to denote participation in the presence of random assignment or, more accurately, 'application to and acceptance into', the training programme, where $d^* = 1$ if an individual is both eligible and assigned to the training (treatment) group and 0 if they are eligible but assigned to the control group, sometimes referred to as eligible non-participants (ENPs). Let r denote randomisation, where $r = 1$ if an individual is randomly assigned to the treatment group and $r = 0$ if an individual is randomly assigned to the control group. Randomisation is implemented at the stage where d^* is revealed.[42]

Assuming random assignment does not alter the probability that individuals seek admission to the programme or their behaviour, that there are no close substitutes for the training programme available and that the sample is sufficiently large, then:

[42] It is possible to design and implement more elaborate random evaluation designs as the participation step can actually be broken down into a number of additional steps, for example: awareness of the programme; realisation of eligibility; application; assessment by administrators; acceptance; and finally training. Randomisation can potentially occur at any stage in this sequence, however the optimal placement depends on the evaluation question being addressed, see Heckman (1992, pp. 222-224).

$$F(y_1 \mid d^* = 1, r = 1) = F(y_1 \mid d^* = 1) = F(y_1 \mid d = 1), \tag{3.3a}$$
$$F(y_2 \mid d^* = 1, r = 0) = F(y_2 \mid d^* = 1) = F(y_2 \mid d = 1), \tag{3.3b}$$

and

$$E(y_1 \mid r = 1) - E(y_2 \mid r = 0) = E(y_1 - y_2 \mid d = 1) \tag{3.4}$$

Conditional on the relevant assumptions, social experiments thus overcome the evaluation problem through the use of *randomised out* non-participants (who would otherwise have participated) to directly obtain the missing counterfactual distribution, $F(y_2 \mid d = 1)$.

Heckman and Smith (1995, p. 89) note that random assignment does not actually eliminate sample selection bias, rather it balances the bias between the participant and non-participant samples. They illustrate this using the common coefficient model: $Y = \alpha + \beta d + u$, where Y is the outcome of interest, α is the mean outcome without training, β is the common effect of participation, and u represents individual characteristics not observed by the administrator. Again using d to represent participation and r to represent randomisation the mean outcome for the treatment and control groups would respectively be

$$E(Y \mid d = 1, r = 1) = \alpha + \beta + E(u \mid d = 1) \tag{3.5a}$$

and $$E(Y \mid d = 1, r = 0) = \alpha + E(u \mid d = 1) \tag{3.5b}$$

The impact of training would hence be captured by the coefficient β, the difference between the two means. The selection problem arises when participation d depends on the unobserved shock u, so that $E(u \mid d) \neq 0$. Randomisation cannot ensure that $E(u \mid d^* = 1) = 0$ or $E(u \mid d = 1) = 0$, that is remove the sample selection bias, rather it balances the bias between the two samples so as it cancels out when calculating the mean impact.

The assumption that random assignment does not alter the probability that individuals seek admission to the programme or their associated behaviour can be formally expressed as $\Pr(d^* = 1) = \Pr(d = 1)$ conditional on the relevant explanatory variables. Proponents of experimental evaluation need to invoke this assumption, or assume it is 'practically true', otherwise the underlying characteristics of participants in the presence of randomisation would differ from the characteristics without randomisation, that is the experiment would be subject to *randomisation bias* (Heckman, 1992, p. 218). The reasonableness of this assumption is questioned in light of indirect evidence and discussed in more detail in section 3.4.

In order to avoid the constraint of the randomisation assumption in applied work however, many experimental evaluators utilise a special case, the dummy endogenous variable model of Heckman (1978). In this model, also called the 'common-effect' model, the impact of the treatment is assumed fixed for all participants, which can be written as:

$$Y_1 \equiv \alpha + Y_2 \tag{3.6}$$

Where α is a constant and it is assumed to be the same for everybody. Following Heckman (1992) the dummy endogenous variable model is able to estimate the mean difference for programme impact (equation 3.4) in the presence of randomisation bias because:

$$\begin{aligned}
&E(y_1|r=1) - E(y_2|r=0) \\
&= E(\alpha + y_2 \mid r=1) - E(y_2 \mid r=0) \\
&= \alpha + E(y_2 \mid d^*=1) - E(y_2 \mid d^*=0) \\
&= \alpha \\
&= E(y_1 - y_2 \mid d^*=1) = E(y_1 - y_2)
\end{aligned}$$

In this case even if randomisation changes the pool of individuals being trained the parameter being estimated remains the same.

Heckman (1992) shows that the unrealistic requirement of treatment outcome homogeneity (fixed α) can be relaxed through application of the random response model while still allowing equation (3.4) to be estimated in certain circumstances. This model allows for the idiosyncratic response of individual participants and is equivalent to the random coefficient model discussed in section 3.5 below. Further assumptions must be imposed to estimate additional parameters of interest, such as the median, making it difficult to address a variety of questions of interest to policy makers without applying econometric procedures with their attendant controversial assumptions (Heckman and Smith, 1995; Ham and LaLonde, 1994; also see section 3.5.4). Analysis of how data from social experiments may be utilised in conjunction with non-experimental methods to establish other parameters of interest can be found in Heckman and Smith (1997).

3.4: Limitations to the Experimental Approach

A well-designed experiment has the potential to solve the evaluation problem and eliminate the difficulty of choosing between alternative estimators but social experimentation can present the evaluator with other problems. Some of these problems are unique to experiments, or more problematic in an experimental context, while others are also common to the non-experimental approach. For example Haveman (1986), in his summary review of social experimentation in the United States between 1965 and 1980, provides a litany of problems and their impact on the reliability of experimental evaluation. These included selectivity problems and the sensitivity of results to the specification of the estimating model two of the key critiques of non-experimental methods. Many of the problems flagged in these earlier experiments have however been addressed in more recent evaluations through better design and the application of theoretical advances (Orr, 1999; Meyer, 1996). There are however, a number of limitations that need to be considered when experimental methods are considered for use. The general limitations of attrition, duration and non-response that are common with non-experimental evaluation are covered in later chapters (for a review of these factors from the experimental perspective see Burtless; 1995, Burtless and Orr, 1986). The focus of this section is on those limitations, which directly affect experimental impact estimates and the administration of social experiments.

One of the key limitations to practical field experiments in the social sciences is the problem of *randomisation bias*. If participation in a training programme is subject to random assignment and this alters the pool and or behaviour of potential participants the experiment is subject to randomisation bias (Heckman, 1992, pp. 218-221). In the case of training programmes, for example, individuals that may have anticipated participating in normal circumstances may alter their decision to apply for a programme if selection into that programme is subject to uncertainty. Risk averse individuals will tend to opt out of the programme and the centres responsible for the randomisation of eligible participants may subsequently find it necessary to screen additional candidates in order to fulfil their requisite quotas and this may result in lower trainee quality. Hence, where participation is open to choice, self-selection bias may still ensue.

Analysis of the Job Training Partnership Act (JTPA), the ongoing primary vehicle for delivery of training to young people and disadvantaged adults in the United States, provides indirect evidence that randomisation bias may be quite serious in the training setting. When the JTPA was introduced in late 1982 it was accompanied by the implementation of the National JTPA Study (NJS), which included a randomly assigned control group allowing experimental evaluation. Doolittle and Traeger (1990) in their report on the implementation of the NJS noted, "Expanded recruitment efforts needed to generate the control group draw in additional applicants who are not identical to the people previously served" (p. 121).

Bjorklund and Regner (1996, pp. 93-94) cite administrative records that show 90 per cent of JTPA centres (service delivery areas) declined to take part in the NJS experiment as additional evidence that randomisation bias is a significant problem in experimental practice. This was in spite of a fourfold increase in compensation offered to the JTPA centres and increased expenditure on information packages for both administrators and participants. Doolittle and Traeger (1990) reported that ethical and public relations difficulties with random assignment and the denial of services to control group members were the concerns most often cited by centres declining to participate in the NJS.[43]

It is however possible that the problem of randomisation bias may be minimal when experimental methods are being used to assess completely new programmes (Bjorklund and Regner 1996, p. 91). Heckman and Smith (1995) point out that this

---

[43] The NJS experiment included just 16 out of more than 600 JTPA centres. A table summarising the relative weight given to different concerns with experimentation expressed by administrators in JTPA centres in the Doolittle and Traeger report is reproduced in Heckman (1992, Table 5.1, p. 220).

argument confuses *demonstrations*, or temporary implementations of new programmes for research purposes, with experiments but they do not address the question of randomisation bias directly. In the demonstration case random assignment can be viewed as simply a means of rationing limited programme resources amongst applicants and thus has the potential to minimise any randomisation bias. In practice however, it may alter the behaviour of both administrators and potential applicants depending on how attractive they perceive the programme to be. Nevertheless in most cases demonstrations do have the distinct advantage of being able to address most ethical objections to experiments, discussed below (Orr, 1999, p.18).

Another factor closely related to randomisation bias is what is generally referred to as the *Hawthorne effect*. This is based on the observation that if participants know that they are in the experiment it may affect their behaviour during the programme.[44] It is possible, for example, that they will then work harder than under normal conditions, which implies that the experiment directly affects the outcome of the programme (Bjorklund and Regner, 1996). It is also possible that the behaviour of programme administrators and bureaucrats may be impacted on by the experiment effectively altering the content of the programme. Bjorklund and Reger (1996, p. 91) label the latter *disruption bias*.[45]

Another problem faced by experimental evaluators is *substitution bias,* which arises when any individuals randomly assigned to the control group gain access to close substitutes for the training programme being evaluated. The controls may gain entry

---

[44] The Hawthorne studies, conducted between 1927 and 1932, identified a positive correlation between the productivity of workers and various experimental manipulations, including the brightness of lights, working hours and managerial leadership. The study largely provided the contemporary rationale for human relations within the organisation. The influence of the Hawthorne study has waned in recent decades, however, following the critical time series analysis by Frank, R. H. and Kaul, J. D. (1978) "The Hawthorne Experiments: First Statistical Interpretation" *American Sociological Review* Vol. 43, 623-643. Frank and Kaul included the 'Great Depression' and instances of managerial discipline, in addition to the original variables, in their analysis and discovered that these factors accounted for much of the observed variation in productivity during the experimental period suggesting that the 'Hawthorne effect' may, to some extent, be illusory.

[45] The importance of these factors on the analysis of the experimental approach can be seen in section 3.3 through the recognition that d^* (application to and acceptance into the programme) is jointly determined by programme administrators and the applicant.

to similar training programmes offered by alternative providers or even the same programme under different funding arrangements. In the presence of substitution bias, control group outcomes no longer correspond to the non-participation state. Heckman and Smith (1996, p. 50) cite work with other co-authors that had access to administrative records on the JTPA indicating that 32 per cent of control group members reported receiving training in the 18 months following random assignment.

The *high costs* often associated with social experiments relative to non-experimental data collection is also often emphasised as a major disadvantage of experiments. Heckman and Smith (1997, p. 4) reported that the first round of the NJS cost approximately $US30 million to administer.[46] Haveman (1986, p. 612) reported a total expenditure of $US1.1 billion on just ten social programmes conducted in the United States between 1965 and 1980. Significantly, the proportion of this expenditure attributable to review and administration costs was a staggering $US450 million. Such constraints clearly impact on the design and effectiveness of the experiment.

However, while cost is clearly a major deterrent to social experimentation in many countries, perhaps the most significant hurdles facing evaluators are due to *ethical objections*. Essentially, the necessity to deny places to eligible would-be participants in social experiments raises an ethical dilemma. This factor currently renders experiments politically infeasible in both the UK and Australia, where the respective governments are committed to universal access to nationally funded programmes. Even in the US where social experimentation is officially sanctioned, ethical considerations were the most frequently cited argument given by administrators of centres declining to participate in the JTPA experiment (Doolittle and Traeger, 1990).[47]

[46] All figures expressed in 1983 dollars.

[47] In contrast, Manski and Garfinkel (1992, p.19) observe that social experiments can, in certain circumstances, actually solve an ethical problem. In cases where programme funds are scarce, and hence placements are limited, then randomisation can provide an equitable means of allocating places.

The ethical dilemma does not entirely preclude experiments from such countries however, as pilot schemes and other such demonstrations can still be administered although establishing the validity of controls would require careful analysis.[48] In their survey of European experiments Bjorklund and Regner (1996, pp. 103-104) cite the *Restart* programme as a rare example of a social experiment in the UK. Although not a training programme, it is useful to briefly review the Restart programme as it not only adopted a number of design innovations to ease some of the objections to experiments but also highlighted some of the practical problems associated with interpreting experimental data.[49]

The evaluation of Restart was commissioned by the Employment Service and analysed by the Policy Studies Institute, an independent research organisation, between March and June 1989. Restart had been in operation on a national scale since 1987 with the dual objective of fostering job search activity through the provision of information about vacancies and other opportunities, and to deter claims for benefits in the absence of eligibility. The programme was targeted at those individuals who had been claiming the unemployment benefit for an unbroken period of six months or longer. Eligible participants were invited for an interview with an officer (the 'Restart counsellor') at the local employment centre. Attendance on a given date was effectively compulsory in the sense that benefits could otherwise be withdrawn. Although interviews typically only lasted 15-25 minutes the aim was to reach a clear outcome, such as referral to a job or training programme. Restart thus became one of the points from which individuals were referred to other active labour market programmes although these programmes were also accessible via other routes. In this sense it is very similar to the Gateway interview in the UK's latest youth training programme, the New Deal.

A random sample of 8,925 eligible participants was identified for the study of which 512 were assigned to a control group and were hence not invited for a Restart

[48] For an informative discussion on evaluation strategies using demonstration studies see Nathan (1991).

[49] White and Lakey (1992) carried out the original analysis on Restart, see also Dolton and O'Neill (1996). Both studies recorded positive findings for the programme.

interview. Appropriate data was compiled on personal characteristics, gender and age, supplemented by information on the Restart interview and outcome. A follow-up survey interview was conducted approximately six months after the scheduled Restart interviews, which was completed by 5,200 respondents. The final sample totalled 4,565 of which 323 were members of the control group (see White and Lakey, 1992, for more detail on the sample construction).

The problem of denying eligible individuals access to the programme affected the design of the experiment in several ways. The size of the control group was relatively small and regionally scattered and those that asked for a Restart interview were given an interview. While these aspects of the design seemed to eliminate ethical objections to the study it does infer a strong probability of substitution bias and impacts on the validity of the control group. The analysis of the data focussed mainly on unemployment duration and exit rates out of unemployment. The researchers concluded the effects of the programme were positive and quite strong as against the low cost of the interview. However, even though the other active programmes were available to the control group it is possible that some of the effects were caused by the programmes in which unemployed persons were placed following the interviews rather than by the interviews *per se* (Bjorklund and Reger, 1996, p. 104).

It would appear that the Restart programme was successful in overcoming ethical and cost limitations to the conduct of a social experiment but causality between treatment (the interview) and effect (reduction in unemployment or duration of unemployment) in Restart is not well defined.[50] Moreover, as the results are subject to substitution and attrition bias, the total effects may not have been elucidated by the evaluation.

In summary, it is clear that social experiments are not yet the panacea many analysts hoped for. Moreover, although experiments of this type can assist evaluation, it is clearly not an optimal form of allocating training resources and given the limitations

---

[50] A caveat is appended to the conclusion in Dolton and O'Neill (1996) raising the question, in light of their positive findings, whether conducting an experiment which permanently effects someone's lifetime prospects is ethically justifiable.

due to practical considerations it is difficult to envisage setting up properly controlled social experiments. Moreover, experimental schemes cannot adjust for voluntary participation in training since the participating group from which controls are drawn remains self selected and of those randomly selected for training some may choose not to participate in training. As a result, what is estimated is the mean impact of training conditional upon participation.

Finally, as observed at the end of section 3.3, only a limited range of questions of interest can be answered using experimental methods within social science settings, indicating a necessary and complementary role for non-experimental methods. This interrelationship is recognised in Ham and LaLonde (1994) and enlarged upon by Heckman and Smith (1995, p. 48) who maintain that

> "Since experiments can answer only a subset of the questions of interest to evaluators, it remains important to build up the stock of basic social science knowledge required to successfully utilize nonexperimental methods both by themselves and as a tool for more extensive analyses of experimental data".

3.5: Non-experimental Methods

Given the limitations to social experimentation discussed above, social scientists rarely have access to true experimental data and evaluators in the main rely on non-experimental data sources.[51] Indeed, as observed by Winship and Mare (1992), if we were to rely solely on observational samples that are free from selection bias we would "...rule out a vast portion of fruitful social research" (p. 328). The critical objective of non-experimental methods is to adjust outcomes to remove any pre-existing differences between participants and the comparison (control) group that would otherwise be mistaken as programme impact (Bell, Orr, Blomquist and Cain, 1995). Thus the non-experimental approach to solving the evaluation problem necessarily requires the evaluator to select a comparison group and to specify an appropriate model of participation and outcome.

---

[51] 'True' experimental data is taken here to mean the type of data often available in laboratory experiments where the appropriate methods can be applied to independently vary treatments and hence isolate spurious channels of causation.

This section begins by looking at alternative comparison groups that might be used in the absence of a randomly assigned control group. Attention then turns to different non-experimental models focusing first on the conventional sample selection model, in which the participation decision is modelled as an underlying latent variable. Commonly applied variants and extensions to the basic model are then examined. To complete the section we explore criteria of interest other than the mean impact of training.

3.5.1: Comparison Groups

Choice of the appropriate comparison group, to act as a counterfactual to participants, can help minimise the differences that lead to sample selection bias. Comparison groups have variously been drawn from eligible non-participants, unsuccessful applicants, programme dropouts, the population at large or even participants themselves. Historically, evaluations routinely adopted participant before-and-after comparison designs with any difference in outcomes being attributed to the programme. Recognition that other influences, such as location or changes in the state of the economy, also impacted on outcomes necessitated the adoption of comparison groups made up of various non-participant populations. This allowed conventional difference-in-differences comparison designs to be used, where before-and-after differences for participants were compared with those of non-participants.

Experiments naturally create *internal* comparison groups through the randomisation of a subset of would-be participants into a control group. Some early non-experimental evaluation studies used internal comparison groups, such as unsuccessful applicants; because they possess many of the same attributes as the participants but for various reasons do not participate. However, growing awareness of sample selection bias, whether due to the participants, non-participants or the administrator, cast doubt on the reliability of this approach. Bell, Orr, Blomquist and Cain (1995) have recently revived some interest in the use of internal comparison groups in a study using applicants that were 'screened out' by programme administrators. Assuming that the unsuccessful applicants were excluded on factors observable to the administrators the researchers applied a 'regression discontinuity model' to evaluate the programme. Their model attempts to control for differences between the participant and comparison groups by using administrators' ratings of applicant potential. Bell et al. assessed the performance of the regression discontinuity model relative to an experimental control group (NJS) and two other internal comparison groups, no-shows and withdrawals (attritors). They found that the impact estimates of their preferred model were more closely aligned to the experimental estimates than the other controls but still suspect.

Most studies use *external* comparison group designs in order to minimise observable sample selection problems. Consisting of a sample of individuals with similar attributes to programme participants, comparison groups are often drawn from national databases, such as the National Child Development Survey (UK) or the Australian Youth Survey. The use of external comparison groups sometimes involves searching for a group of individuals who are statistically matched to programme participants. Matching involves the pairing of participant and comparison group members either according to socio-economic characteristics (cell matching) and/or according to eligibility criteria of the programme (eligibility match). The hope is that, by matching in terms of observed attributes, the two groups will also be similar in terms of their unobserved characteristics (Riddell, 1991, p. 57). In effect, the matching estimates are designed to approximate the virtues of randomisation and the difference in mean outcomes of the two groups is then the desired measure of programme impact.[52] Recently, Dehejia and Wahba (1998) and Heckman, Ichimura and Todd (1997) have adopted the propensity score method, developed in the field of statistics by Rosenbaum and Rubin, (1983) to improve matching in evaluation studies. The propensity score summarises the information in a set of observable variables into a single index function. The

[52] A distinction can be drawn between statistical and econometric matching as the latter "...denotes the standard behavioural modelling techniques and specification tests that use observed characteristics as regressors, without necessarily restricting the composition of the estimation sample through the use of matching" (Friedlander, Greenberg and Robins, 1997, p. 1818). Both approaches, however, adjust the outcome estimates for a given set of observable covariates (regressors) and differ only in the way in which they specify the functional relationship between observables and the relevant outcome variable. Alternative matching techniques, such as cell matching, eligibility matching and individual nearest neighbour matching, are discussed in Barnow (1987) and Dickinson Johnson and West (1987).

participation and comparison group observations with similar propensity scores, in other words similar predictions of being in the treatment group, are considered good matches for each other. Dehejia and Wahba (1998) argue that this method can yield accurate impact estimates by paring down the large pool of potential controls to the relevant comparison group without using outcome information. This greatly reduces data demands and they find that the matched group is reasonably successful in replicating the characteristics of an experimental control group generated from the same data. Ironically, Dehejia and Wahba use the same dataset employed by LaLonde (1986) who concluded comparison group methodology had failed to provide reliable impact estimates and that experimental designs were preferable. LaLonde's analysis of non-experimental estimators is discussed below in section 3.6.

Heckman et al. (1997) extend the matching framework of Rosenbaum and Rubin (1983) by adopting a local linear matching method, which includes a linear term in the probability of participation when constructing matches. Their analysis suggests that mismatching and inaccurate weighting of data is numerically more important than selection bias due to unobservable differences. Matching proved more effective in recovering the parameters of interest when both groups are drawn from the same local labour market, and when the same questionnaire was administered to both groups in the sample. The latter is designed to ensure outcome and personal characteristics are measured in the same way. Contrasting the characteristics and outcomes of the matched data with JTPA experimental data (the NJS control group), Heckman et al. conclude that matching helps reduce but may not eliminate selection bias. Clearly, the formation of exact matches is computationally demanding and may result in relatively small samples.

There is still considerable disagreement on appropriate comparison group design, nevertheless it is clearly important that comprehensive background information should be collected on all sample members to support statistical and econometric controls for observable pre-programme differences between participant and control group members. This will help eliminate any selection bias where it is based on observable heterogeneity. Of more concern to analysts, however, has been the identification and mitigation of selection bias where it arises from unobservable characteristics of sample members.

3.5.2: Selectivity Models

In the presence of unobservable heterogeneity, the evaluator must not only specify an appropriate model relating the outcome of interest to the programme but also one that corrects for any underlying differences between the comparison and participant groups. Heckman and Smith (1996) aptly describe the process of constructing the relevant counterfactuals within the non-experimental approach as "...subjective thought experiments, in which assumptions replace missing data" (p. 52). This is because, in the absence of random assignment to a training programme, it is necessary to model the decision (selection) rules governing participation in that programme. Specific statistical assumptions concerning the relationship between the participation decision and the individual's future employment or earnings provide a way to equalise the starting point for the participant and comparison groups when estimating the impact of the programme. If these assumptions hold, the resulting impact estimates can then be considered unbiased.

Heckman and Robb (1985), focusing on the issue of identification, carried out a comprehensive survey of non-experimental procedures that can be implemented for alternative types of data, cross-section, repeated cross-section, and longitudinal. They clarified the variety of different assumptions that might be made to make estimation of the impact of training on earnings feasible in the presence of sample selection bias. The key argument of the paper was, however, that just one assumption suffices to identify the mean impact of training on earnings. Since different assumptions appeal to different people, econometric estimators were subsequently interpreted as arbitrary and unreliable by many observers (Ashenfelter and Card, 1985; LaLonde, 1986; Burtless and Orr, 1986). In contrast, Heckman and Smith (1995) argue that the key element for improving estimation rests on the better data.

Heckman and Robb (1985) identified three main approaches to overcoming the evaluation problem. One of these exploits the availability of longitudinal data by utilising both prior and post-programme observations on individual participants. The motivation for this approach rests mainly on the assumption that pre-programme observations of earnings may provide a good measure of unobservable determinants of post-programme earnings.[53] The limited labour market experience of most youth, however, generally restricts both data availability and the applicability of this type of model for the purposes of evaluating youth training programmes. This limitation is exacerbated by the likelihood that better impact estimates of programme effect using this approach rest on the quantity of data (Moffitt, 1991, p. 300).

Alternatively, we can impose parametric distributional assumptions on the participation equation or utilise an identifying variable(s) that determines participation but does not affect the outcome. The difficulty with the latter is finding a suitable variable as discussed below. The main focus of this sub-section is therefore on methods that rely on distributional assumptions for model identification. The availability of broad survey data containing information on both programme participants and non-participants, such as the Scottish Young Person's Survey, together with the relative paucity of prior labour market experience of youth training participants lends itself to such an approach. First, the conventional (Heckman) two-step model, which rests on the assumption of normality,[54] is outlined in some detail. This model has been extensively applied in a wide array of evaluation studies. It has also been modified and extended in an impressive variety of ways and here we consider a number of varieties that are applicable in the context of evaluating youth training programmes. Barnow, Cain and Goldberger (1981) incorporated the training choice in the outcome equation. Lee (1982, 1983) allows for a more general form of the model allowing for different distributional transformations and an extension to polychotomous choice. The treatment model can also be generalised to the

---

[53] Card and Sullivan (1985) provide a good example of this methodology using longitudinal data on CETA participants and data on a comparable group of non-participants from the Current Population Survey (USA). They are able to construct the earnings history of individuals in a number of pre and post-training periods and hence are able to compare the trainee's predicted post-programme earnings, based on a time-series determination of earnings in the comparison group, with their actual post-programme earnings.

[54] Just as, for example, certain inherent biological traits are generally thought of as being distributed normally in the population at large (Moffitt, 1991, p. 304).

endogenous switching model, which allows the effect of the independent variable to vary across treatments. The vast majority of non-experimental evaluation of training programmes relies on these basic approaches (see chapter 4). Although our focus is on evaluation the same basic methodology can also be used in a variety of different applications in the wider field of labour economics, see Heckman and MaCurdy (1986) for a survey.

The Standard Selectivity Model

To draw on the essential aspects of the econometric approach and assist clarity we again assume two potential states for each individual but now adopt more familiar linear equations to represent the relevant processes. Our initial focus is on post-programme earnings as the outcome of interest, denoted by y, post-programme employment is considered later in this section. Following Maddala (1983) and Heckman (1979), the typical two-equation selection model can then be written as:

$$y_{1i} = X_i\beta_1 + u_{1i} \quad \text{(for participants)} \tag{3.7a}$$
$$y_{2i} = X_i\beta_2 + u_{2i} \quad \text{(for nonparticipants)} \tag{3.7b}$$

where for each individual i, the outcome of interest y, is controlled for a vector of observable characteristics (such as: age, gender, education and experience) denoted by X, with the associated vector of parameters to be estimated β, while any unobservable factors affecting y are captured in the random error term u. The error term is assumed normally distributed with mean zero and variance σ_u^2.

The evaluator does not, however directly observe the decision to participate in training, only whether individuals are in training or not. Hence we employ a latent variable equation to represent the decision process for individual i with given personal characteristics and environmental conditions, denoted by the vector Z. This process can be expressed as:

$$I_i^* = Z_i\gamma - \varepsilon_i \quad \text{where:} \quad \begin{array}{l} I_i = 1 \text{ iff } I_i^* > 0 \\ I_i = 0 \text{ iff } I_i^* \le 0 \end{array} \tag{3.7c}$$

where I_i is the dichotomous realisation of the participation process, ε is normally distributed with mean zero and unit variance. The observed y_i is then defined as:

$$y_i = y_{1i} \text{ iff } I_i = 1$$
$$y_i = y_{2i} \text{ iff } I_i = 0 \tag{3.7d}$$

$$\text{Cov}(u_{1i}, u_{2i}, \varepsilon_i) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1\varepsilon} \\ \sigma_{12} & \sigma_{22} & \sigma_{2\varepsilon} \\ \sigma_{1\varepsilon} & \sigma_{2\varepsilon} & 1 \end{bmatrix}$$

The dependent variable y_{1i} conditional on X_i and Z_i has a well-defined marginal distribution but it is not observed unless $I_i^* > 0$. In other words the observed y's of y_{1i} are censored [55] and

$$y_{1i} = X_i\beta_1 + u_{1i} \text{ iff } Z_i\gamma > \varepsilon_i$$

In the terminology adopted by Heckman and Robb (1985, 1986), selection bias occurs because of the stochastic dependence between u_{1i} and ε_i, that is selection on *unobservables*. For each participant with characteristics X_i and Z_i we want to compare the outcome of being in the programme y_{1i} with the expected potential outcome if they did not partake in the programme, that is the counterfactual $E(y_{2i}|I_i = 1)$.[56] Under the normality assumption the gross benefit for individual participant i can thus be written as:

$$\begin{aligned} y_{1i} - E(y_{2i} \mid I_i = 1) &= y_{1i} - X_i\beta_2 + E(u_{2i}|I_i = 1) \\ &= y_{1i} - X_i\beta_2 + \sigma_{2\varepsilon}\frac{\phi(Z_i\gamma)}{\Phi(Z_i\gamma)} \end{aligned} \tag{3.8}$$

Where $\sigma_{2\varepsilon}$ is the covariance between u_{2i} and ε_i and the ratio on the right-hand side is the Inverse Mill's Ratio or non-selection hazard function, usually symbolised as λ_i

[55] The modern literature on the estimation of models for censored data can be traced to Tobin's (1958) paper, which analysed the relationship between the ratio of expenditure on durable goods to total disposable income. The model Tobin introduced has become known as the Tobit and is now widely employed in the literature with many generalisations and a wide range of estimation methods. Amemiya (1985) set out to classify the diverse range of models and assigned Heckman's two-step estimator as a 'Type 2 Tobit', see also Breen (1996) and Maddala (1983). An alternative to the selectivity model is the closely related Double Hurdle model developed by Cragg (1971). This model defines the participation (observability) rule in a slightly different fashion requiring both the indicator variable I^* and the outcome variable y^* to be positive before an observation is recorded for the dependent variable. For further discussion on Cragg's model see Breen 1996.

[56] $E(y_{2i}|I_i = 1)$ is a counterfactual as the condition $(I_i = 1)$ precludes the variable over which the expectation is taken and there is no way to estimate it directly using data.

(i.e. the ratio of the standard normal population density function ϕ to the cumulative distribution function Φ evaluated at $Z_i\gamma$).[57] The sum of these two terms corrects for the selection bias and hence a viable test for selection bias is if $\sigma_{2\varepsilon} = 0$ (Melino, 1982). Since the components of λ_i depend only on the participation process, it can be estimated from the results of Probit analysis of the participation decision, the first step in the Heckman two-step method.

In other words, to fully evaluate the benefit of the training to the participants from a cost-benefit standpoint we thus need to first estimate the conditional expectation of u_{2i} and then sum equation (3.8) for all participants. If, as we hypothesised in section 3.2, those individuals who have a comparative advantage in training are the most likely to participate, then we would expect that the programme will produce greater benefits than under random assignment. The expected gross benefit for participant i can be determined by estimating the difference between their expected earnings from participation and their expected earnings had they not participated (the counterfactual), that is:

$$E(y_{1i} \mid I_i = 1) - E(y_{2i} \mid I_i = 1) = X_i(\beta_1 - \beta_2) + (\sigma_{2\varepsilon} - \sigma_{1\varepsilon})\frac{\phi(Z_i\gamma)}{\Phi(Z_i\gamma)} \tag{3.9}$$

Given that the γ parameters on Z_i can be estimated from the maximum likelihood of participation using Probit estimation, it is possible to estimate equation (3.9) using OLS, that is the second step in the Heckman two-step method. If the self-selection described is present then the covariance terms $\sigma_{2\varepsilon} - \sigma_{1\varepsilon}$ will be greater than zero. As observed above, the product of this and the Mill's ratio will remove the selectivity bias and allow consistent estimation using OLS by accounting for the 'omitted variable' λ. By including λ as an additional variable we are hence effectively

---

[57] Strictly speaking the Mill's ratio is evaluated at $Z_i\gamma/\sigma_\varepsilon$ however, under the standard normal it is generally assumed that $\sigma_\varepsilon = 1$, because γ and σ_ε are not separately identifiable (Breen, 1996, p. 36). The expression used for the Mill's ratio in equation (3.8), $\phi(Z_i\gamma)/\Phi(Z_i\gamma)$ or the non-selection hazard based on the non-selection equation, follows Maddala (1984, p. 261). This is equivalent to $\phi(Z_i\gamma)/1 - \Phi(Z_i\gamma)$ or the non-selection hazard based on the selection equation, as used by Heckman (1979, p. 156).

controlling for the difference in outcomes due to self-selected participants (those with a comparative advantage in training). It does not, however, account for heteroskedasticity and hence the standard errors need to be corrected. Moreover, as we are only using an estimate of λ, the standard errors for the β coefficients need to take into account any error in estimating λ (Breen, 1996, p. 38). The following procedure, based on Heckman (1979, p.157), can be used:

Let $\delta_i = -\lambda_i(I_i + \lambda_i)$ where $I_i = Z_i\gamma$ (from the Probit equation)

Then let s_u be the uncorrected estimate of σ_u from the outcome regression and let S be the sum of squared deviations from this regression

$$S = \sum_{I=1} (y_i - \hat{y}_i)^2 \qquad \text{for all } I_i = 1$$

The correct asymptotic estimate of σ_u is then given by

$$\hat{\sigma}_u = \frac{1}{n}\left[S - \hat{\sigma}_{u\varepsilon}^2 \sum_{i=1}^{n} \delta_i\right]^{\frac{1}{2}} \tag{3.10}$$

Where n is the number of observations for which $I = 1$ and $\hat{\sigma}_{u\varepsilon}$ is the estimated coefficient for λ. It is possible to obtain a valid covariance matrix estimate although the calculation is cumbersome and the estimated covariance matrix is not always positive definite (Greene, 1994, pp. 774-745; Breen, 1996, p. 38-39). Such corrections are, however, routinely included in most relevant computer packages such as LIMDEP (Green, 1995) or STATA.

Alternatively, the sample selection model can be estimated using the maximum likelihood method. Given the observation type is determined by the underlying decision process indicated by the value of I_i, and the log likelihood function is made up of expressions for the probability of observing each value of y, it can then be written as

$$L = \sum_{I=0} \log(\Pr(I_i = 0)) + \sum_{I=1} \log(\Pr(I_i = 1) f(y_{1i} | I_i = 1) \tag{3.11a}$$

The first term in equation (3.11a) is the summation, over all observations for which I_i = 0, of the logarithms of the probability that I_i = 0. The second term is the summation over all observations which $I_i = 1$ multiplied by the density of y_{1i} conditional on $I_i = 1$. The second term can also be written as

$$\sum_{I=1} \log(\Pr(I_i = 1 | y_{1i}) f(y_{1i})$$

Where $f(y_{1i})$ is the unconditional density of y_{1i}, which is the normal density with conditional mean $X_i\beta$ and variance σ^2. Since u and ε are bivariate normal, with correlation coefficient ρ we can write

$$I_i^* = Z_i\gamma + \rho(\frac{1}{\sigma}(y_{1i} - X_i\beta)) + \varepsilon$$

It follows that

$$\Pr(I_i = 1) = \Phi \frac{Z_i\gamma + \rho\left(\frac{(y_{1i} - X_i\beta)}{\sigma}\right)}{(1-\rho^2)^{1/2}}$$

Since $y_i = y_{1i}$ when $I_i = 1$. Thus the log likelihood function can be written as:

$$L = \sum_{I=0} \log(\Phi(-Z_i\gamma) + \sum_{I=1} \log\frac{1}{\sigma}(y_i - X_i\beta)^2 + \sum_{I=1} \log\Phi\left[\frac{Z_i\gamma + \rho\left(\frac{y_i - X_i\beta}{\sigma_u}\right)}{(1-\rho^2)^{1/2}}\right] \tag{3.11b}$$

Maximum likelihood estimates can then be obtained in the standard method by maximising equation (3.11b), see also Amemiya (1984, pp. 31-32). Note that when $\rho = 0$ this equation reduces to two parts: a standard Probit for the probability of being selected and an OLS for the expected value of y in the selected sub-sample. In this special case, as there are no common parameters in either part, they can be estimated separately showing that, if there is no residual correlation between u and ε, the standard OLS approach is adequate. Moreover, this illustrates that the difficulty underlying the problem of sample selection is due not to censoring alone but to non-random selection of observations with respect to y (Breen, 1996, p. 40).

The two-step method is, however, used more frequently for the estimation of sample selection models. As observed by Davidson and MacKinnon (1993, p. 544) this rests primarily on its relatively simple computational cost in comparison to maximum likelihood estimation. On the other hand, Breen (op cit, p.71) states that "Given the availability of ML methods [in modern econometric packages], there is little to be said in favor of using the original two-step approach". The relative performance of the two methods when using artificially generated data is discussed below in section 3.7.

The Treatment Effects Model

Some studies employ a model with only a single outcome equation, sometimes referred to as the treatment effects model (Greene, 1993, p. 713) of the form:

$$y_i = X_i\beta + I_i\alpha + u_i \tag{3.12a}$$

This is analogous to the standard selectivity model, excepting that the participation dummy I appears directly in the outcome equation. Hence, in the absence of self-selection, the estimate of α would measure the effectiveness of the training programme for participants. In the presence of sample selection, however, we again need to model participation as before:

$$I_i^* = Z_i\gamma - \varepsilon_i \quad \text{where:} \quad \begin{array}{l} I_i = 1 \text{ iff } I_i^* > 0 \\ I_i = 0 \text{ iff } I_i^* \leq 0 \end{array} \tag{3.12b}$$

As before, sample selection arises if u_i and ε_i, are correlated. Barnow et al., (1981) suggest two methods for estimating this model. It can be estimated in the same fashion as the standard model but using the entire sample of data. Alternatively, estimate the model by non-linear two-stage least squares using as the instrumental variable for I the predicted probabilities from the Probit equation. While such an approach does allow the effect of the programme to be observed readily it does constrain the coefficients of the equations for both participants and non-participants to be the same and the variable I cannot be treated as exogenous if the decision to participate is based on self-selection. If the variable I is endogenous, equation (3.11) must be estimated using instrumental variable techniques, discussed below.

The Binary Logit Selection Model (Lee model)

The Probit-OLS approach in the standard model requires normality of the selection equation error. Implicitly, the standard approach measures the effect of the treatment by deviations from normality of the distribution y_i within the sample of participants. However, the realism of the normality assumption has been questioned by amongst others Olsen (1980, 1982) and Golberger (1981) who demonstrated that existing estimators for censored regression (Tobit and related) models are not robust to departures from normality. Lee (1983) describes a reformulation of the selection model that allows more general specifications of the selection and outcome equations. Consider the following two equation censored model:

$$y_1 = X_i\beta + u_i \qquad (3.13a)$$
$$y^* = Z_i\gamma - \varepsilon_i \qquad (3.13b)$$

Where X and Z are exogenous variables and the selection or 'choice' rule is defined by equation (3.13b). As before the dependent variable y^* is not directly observed by the evaluator but has a dichotomous realisation I that is related to y^* by the following process

$$I = 1 \text{ iff } y^* \geq 0, \quad \text{otherwise} \quad I = 0$$

Hence the dependent variable y_1 is only observable if $y^* \geq 0$, that is the observed distribution of y_1 is censored although, conditional on X and Z, it has a well-defined marginal distribution. Let $F(\varepsilon)$ and $G(u)$ be the distribution functions for ε and u and assume these are known or completely specified. Let $\Phi(\cdot)$ be the standard normal distribution function and let $B(\cdot, \cdot; \rho)$ be the bivariate normal distribution with zero means, unit variances and correlation coefficient ρ. Given the completely specified marginal distributions of ε and u, each of them can be transformed to the standard normal random variable $N(0, 1)$. Let

$$\varepsilon^* = J_1(\varepsilon) \equiv \Phi^{-1}[F(\varepsilon)] \qquad (3.14a)$$
$$u^* = J_2(u) \equiv \Phi^{-1}[G(u)] \qquad (3.14b)$$

Both the transformed random variables ε^* and u^* are standard normal $N(0, 1)$ variables with zero means and unit variances. In this case, the transformations J_1 and

J_2 involve the inverse of the standard normal distribution function although a rich variety of specified transformations are possible (for examples see Lee, 1982). A bivariate distribution having the marginal distributions $F(\varepsilon)$ and $G(u)$ can be specified as

$$H(\varepsilon, u; \rho) = B[J_1(\varepsilon), J_2(u); \rho]. \tag{3.15}$$

The bivariate distribution of (ε, u) is derived by assuming that the transformed variables ε^* and u^* are jointly normally distributed with zero means, unit variances, and correlation coefficient ρ. When $\rho = 0$, it corresponds to statistical independence of ε and u. When the marginal distributions of ε and u are normally distributed, the above bivariate distribution will then be a bivariate normal distribution. Let (y_i, x_i, z_i, I_i), $i = 1,\ldots, N$, be the random samples. The log likelihood function based on this specification is thus:

$$L = \sum_{i=1}^{N} \{ I_i \log g\left(\frac{y_i - X_i\beta}{\sigma}\right) + I_i \log \Phi\left(\frac{J_1(Z_i\gamma) - \rho J_2 \dfrac{(y_i - X_i\beta)}{\sigma}}{\sqrt{1-\rho^2}}\right)$$
$$- I_i \log\sigma + (1 - I_i)\log(1 - F(Z_i\gamma)\} \tag{3.16}$$

Where g is the density function of u. Maximum likelihood estimates can then be obtained in the standard method by maximising equation (3.16).

Alternatively two stage estimation methods can be used to obtain initial consistent estimates, as illustrated by Lee (1983, p. 508-509) who demonstrates the special case of the bivariate normal distribution with the only assumption necessary being that u are $N(0, \sigma^2)$, whereas the distribution of ε is arbitrary. Lee's model can, however, start from any appropriate bivariate distribution and after transforming the random variables it is possible to construct a bivariate distribution with given margins F and G. The first step involves modelling the selection process, based on whatever assumption for the error is valid, and computing the predicted probabilities of participation. Estimation in the second step can then proceed as in the standard model after substituting $J_1(Z_i\gamma)$ for $Z_i\gamma$. For example, if ε follows the normal distribution, then we have the standard Probit-OLS method of correction for selection bias. If $F(\varepsilon)$ is the logistic distribution, then we have the Logit-OLS method of correction for selection bias. The methods that can be used to estimate (3.12) can thus be very general. Lee (1983, pp. 509-510) demonstrates the approach when the choice equation is assumed to be student t distribution. Maddala (1983, p. 274) presents a more general specification that can be used for other transformations. The most common form of application of the general model involves the utilisation of a Logit for the selection equation, that is the binary Logit selection model. This model can also be readily extended to incorporate multiple options (see the discussion on polychotomous choice models below).

The Switching Regression Model

The endogenous switching model can be applied when the effects of the independent variable are thought to vary across treatments. Following Maddala (1983) the model can be defined in the same framework as the standard model (equations 3.7a-3.7d), that is:

$$y_{1i} = X_{1i}\beta_1 + u_{1i} \qquad (3.17a)$$

$$y_{2i} = X_{2i}\beta_2 + u_{2i} \qquad (3.17b)$$

$$I_i^* = Z_i\gamma - \varepsilon_i \quad \text{where:} \quad \begin{array}{l} I_i = 1 \text{ iff } I_i^* > 0 \\ I_i = 0 \text{ iff } I_i^* \le 0 \end{array} \qquad (3.17c)$$

The observed y_i is defined as

$$\begin{aligned} y_i &= y_{1i} \text{ iff } I_i = 1 \\ y_i &= y_{2i} \text{ iff } I_i = 0 \\ (u_1, u_2, \varepsilon)' &\sim N(0, \Sigma) \end{aligned} \qquad (3.17d)$$

with $\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1\varepsilon} \\ \sigma_{12} & \sigma_{22} & \sigma_{2\varepsilon} \\ \sigma_{1\varepsilon} & \sigma_{2\varepsilon} & 1 \end{bmatrix}$

If the model has $\sigma_{1\varepsilon} = \sigma_{2\varepsilon} = 0$ the switching is considered exogenous otherwise it is endogenous. Equations (3.17a) and (3.17b) define the marginal distributions of y_{1i} and y_{2i}. It is then possible to derive the conditional distributions $f(y_{1i}|I_i = 1)$ and $f(y_{2i}|I_i = 0)$ and:

$$f(u_1 | I = 1) = \int_{-\infty}^{Z\gamma} f(u_1 \varepsilon) d\varepsilon \,/\, \Phi(Z\gamma)$$

and writing $f(u_1 \varepsilon) = f(u) \cdot f(\varepsilon | u)$, we can write:

$$f(y_1 | I = 1) = \left[\Phi(Z\gamma)\right]^{-1} \sigma_{11}^{-1/2} \phi\left[\sigma_{11}^{-1/2}(y_1 - X_1\beta_1)\right]$$
$$\times \Phi\left\{\left(1 - \frac{\sigma_{1\varepsilon}^2}{\sigma_{11}}\right)^{-1/2}\left[Z\gamma - \frac{\sigma_{1\varepsilon}}{\sigma_{11}}(y_1 - X_1\beta_1)\right]\right\} \tag{3.18a}$$

Similarly,

$$f(y_2 | I = 0) = \left[1 - \Phi(Z\gamma)\right]^{-1} \sigma_{22}^{-1/2} \phi\left[\sigma_{22}^{-1/2}(y_2 - X_2\beta_2)\right]$$
$$\times \left(1 - \Phi\left\{\left(1 - \frac{\sigma_{2\varepsilon}^2}{\sigma_{22}}\right)^{-1/2}\left[Z\gamma - \frac{\sigma_{2\varepsilon}}{\sigma_{22}}(y_2 - X_2\beta_2)\right]\right\}\right) \tag{3.18b}$$

Greene (1995, p. 668) notes that, in contrast to the standard model, the separation in the endogenous switching model is stronger, in the sense that at any given observation only one regime is observed, that is only one of X_{1i} or X_{2i} is observed with y_i. The model can thus be estimated one equation at a time, either by two-stage least squares or maximum likelihood but, in this setting, both these approaches ignore some information. The preferred estimation method is then a full information maximum likelihood that jointly estimates $\beta_1, \beta_2, \sigma_{11}, \sigma_{22}$, and ρ.

Instrumental Variables

The method of instrumental variables is increasingly used in evaluation studies having been proposed by Heckman and Robb (1985, 1986). Successful estimation using this method rests on an identifying variable(s) that determines participation but has no direct affect on the outcome. Provided that a legitimate instrumental variable can be found there are a variety of estimation techniques available however, following Heckman and Smith (1996), we can illustrate the concept by adopting a simplified treatment effects model of the form

$$y = X\beta + d\alpha + u \tag{3.19}$$

Participation is indicated by the dummy variable *d*, where *d* = *1* denotes participation and $d = 0$ otherwise. While *u* is an error term with mean zero, so that $E(u) = 0$ and α is assumed the same for everyone. In this model, correlation between *d* and *u* gives rise to the evaluation problem. For example, if more motivated individuals, with corresponding high values of *u* in the absence of the training programme, were to apply for training in disproportionate numbers then $E(u|d = 1) \neq 0$.

An instrumental variable *Z* needs to satisfy two conditions. First, that it is uncorrelated with *u,* so that $E(u|Z) = 0$ and second, that it is correlated with the participation decision *d,* so that $E(d|Z) = \Pr(d = 1|Z)$ is a nontrivial function of *Z* taking on at least two distinct values for two distinct values of *Z*. These properties can then be applied to the outcome equation and taking expectations conditional on *Z*, we have:

$$E(Y \mid Z) = \beta + \alpha \Pr(d = 1 \mid Z)$$

Allow Z_1 and Z_2 to be two distinct values of the instrumental variable *Z*, such that, $\Pr(d = 1|Z_1) \neq \Pr(d = 1|Z_2)$. The relevant conditional expectations are then

$$E(Y \mid Z_1) = \beta + \alpha \Pr(d = 1 \mid Z_1)$$

and $$E(Y \mid Z_2) = \beta + \alpha \Pr(d = 1 \mid Z_2)$$

If we then take the difference between these two equations and solve for α we obtain

$$\alpha = \frac{E(Y|Z_1) - E(Y|Z_2)}{\Pr(d = 1|Z_1) - \Pr(d = 1|Z_2)} \tag{3.20}$$

The key advantage of the instrumental variable estimator is that it removes any need to invoke distributional assumptions. A major limitation to this approach, however, is the difficulty in finding a credible instrument. One possibility arises when a programme is not universally available due to different political or bureaucratic priorities (Moffitt, 1991). If a random sample of the population of interest were taken from a neighbourhood, city or state where the programme is available, along with one from where it is not, then a comparison of the mean values in the two would form a valid estimate of programme impact α. The variable *Z* in this case should be thought of as a dummy variable equal to 1 in the neighbourhood with the programme and zero in the other. This approach, sometimes termed a *natural experiment* (Meyer, 1996), is similar to an experiment in that the probability of having the treatment available is random with respect to the outcome variable under study as a result of natural variation. As such this form of cross-region evaluation underlies demonstration studies and pilot programmes. If the cross-state variability were due to factors other than political or bureaucratic decisions but on relative labour market conditions it would cease to be a valid instrument (see also Heckman and Smith, 1996, p. 67). Moreover, in the context of universal programme availability, such as in Australia and the UK, the cross-region approach is limited. See also the discussion on comparisons with experimental data in section 3.7 below.

A wide variety of possible instruments have been proposed in the evaluation literature, these may include individual rather than geographical characteristics. For example, Willis and Rosen (1980), who investigated the returns to education, used father's education, mother's work experience, religion and number of siblings, while Card (1984) used the proximity (nearby or not) of a college. Dearden and Heath (1996), investigating the effects of introducing an education allowance payable to relatively disadvantaged 16 and 17 year olds who remain in school in Australia, use the number of siblings aged over 15 as an instrument. In terms of identifying an 'ideal' instrument, Moffitt (1991) notes that it " ...is variation in the availability, rather than the actual receipt, of treatment across the population that is more likely to provide a legitimate *Z*..." (p. 297). Although discussing healthcare Moffitt's observation is equally likely to apply in training and other ALMP evaluation.

It is, however, difficult to choose a legitimate instrumental variable, indeed even seemingly ideal instruments can prove controversial. Angrist (1990) suggested the use of the lottery draft as an ideal instrumental variable for identifying the effect of military service on earnings in the USA. The 1969 draft lottery randomly assigned different priority numbers (*Z*) to individuals of different birth date. The higher the number, the less likely an individual was to be drafted. This would appear to be an ideal instrument as it determines participation but has no direct affect on subsequent earnings. However, Heckman and Smith (1996, pp. 66-67) argue that it is a poor instrument because individuals can still choose to serve in the military if they anticipate gains. Moreover, individuals with a high Z, and hence a low chance of being drafted, are therefore likely to be more attractive to employers planning to invest in their workers. Heckman and Smith (1996, p.66) conclude that in any application of instrumental variables the necessity remains for a "...behavioural assumption be made about how persons make their decisions about programme participation".

Employment Probit

When the outcome of interest is post-programme employment, analysts typically adopt a probabilistic framework and estimate the impact directly from a model such as the Probit. In other words, the outcome of interest is simplified to the probability of being in post-programme employment, y, and the appropriate model can be expressed as

$$y_i^* = X_i\beta - u_i \qquad \text{where:} \quad \begin{array}{l} y_i = 1 \text{ iff } y_i^* > 0 \\ y_i = 0 \text{ otherwise} \end{array} \tag{3.21}$$

Where y_i is a dichotomous variable denoting employment ($y_i = 1$) or unemployment $y_i = 0$) expressed in terms of the indicator variable, y^*. X is a vector of observable characteristics used to describe the young person, with the associated vector of parameters to be estimated β, including a dummy variable to indicate participation in the relevant training programme, while unobservable factors affecting y are captured in the random error term u. The error term is again assumed normally distributed with mean zero and variance σ_u^2. Whether a young person is in employment at a subsequent reference date can hence be expressed in terms of whether the value of the indicator variable exceeds a certain critical value, that is:

$$\begin{aligned} \Pr(y_i = 1) = \Pr(y_i^* > 0) &= \Pr\left(u_i > -\left(X_i\beta\right)\right) \\ &= 1 - \Phi\left(-X_i\beta\right) \end{aligned} \tag{3.22}$$

Where Φ is the normal cumulative distribution function. An estimate of the likelihood, taken over the sample N, of observing those young people variously in employment or not in employment can then be written as:

$$L = \prod_{y_i=0} \Phi(-X_i\beta) \prod_{y_i=1} [1 - \Phi(-X_i\beta + d_i\delta)] \tag{3.23}$$

The influence of any particular descriptive variable, X_k, on the probability of being in employment can then be obtained by taking the first deriviative of equation (3.23) with respect to X_k, that is:

$$\frac{\partial \Pr(y_i = 1)}{\partial X_k} = \beta_k \phi(X_i\beta) \tag{3.24}$$

Where ϕ is the standard normal density function. Estimation of equation (3.24) at the sample mean of the descriptive variables allows the estimated coefficients from the Probit maximum likelihood estimation to be interpreted in a similar fashion to that of OLS coefficients.

3.5.3: Extensions to the Basic Models

A number of factors impact on the fit of the model to the programme being evaluated including, our knowledge of the participation decision making process, the availability of alternatives to training, the quality and design of the requisite dataset, and relevant outcome of interest. For example, in the standard selectivity model participation is depicted as a simple dichotomous decision of whether to participate or not. However, in practice training schemes often provide a variety of different training structures including support for private sector programmes of various kinds such as apprenticeships. The returns to training under different paths within the programme may vary requiring more complex evaluation. Different school-to-work transition routes, such as further or higher education may also impact on the decision to participate in training. In such cases a *polychotomous-choice* model may be more appropriate.

The dynamic nature of the youth labour market may also dictate model design. For example, it is not uncommon for young people to change their labour market status relatively frequently, which raises a number of econometric issues (Bradley, 1995, p. 46). Hence, when the outcome of interest is employment, the choice of the post-programme period at which labour market status of the participants is observed naturally impacts on the probability of observing them in employment. This choice has tended to be arbitrary in most evaluation research, dictated to largely by the relevant follow-up period in the dataset used. Similarly, the nature of government assisted training programmes often result in young people moving in and out of the programme leading to significant non-completion rate. To account for the more dynamic nature of the youth labour market it may be useful to employ a *duration* model. This class of models allows the evaluator to assess issues such as the time it takes a participant to gain useful employment or the duration of subsequent employment spells. Ham and LaLonde (1994) also support the use of duration models for the evaluation of training programme because of the interaction between earnings and employment.

The remainder of this sub-section outlines the polychotomous choice and duration models in more detail, as these approaches have both been applied in the training evaluation literature in Australia and the UK. A number of other more elaborate extensions to the sample selection model are feasible, such as accounting for sequential or joint decision processes (see Maddala, 1983, pp. 278-283 or Breen, 1995, pp. 50-54). Naturally the relevance of the model is dependent on the, type of programme being evaluated, data availability, institutional setting and the question being addressed. For example, Keane and Moffitt (1998) examine the increasingly common situation of individuals simultaneously participating in more than one welfare programme and the resultant impact of this phenomenon on labour supply.

Polychotomous-choice Models

Following Lee (1982) a polychotomous-choice model with sample selection bias, allowing M categories with one potential regression outcome in each path, may be represented as

$$\begin{aligned} y_{si} &= X_{si}\beta_s + u_{si} \qquad (s = 1,2,\ldots,M) \\ I^*_{si} &= Z_{si}\gamma + \eta_{si} \qquad (i = 1,2,\ldots,N) \end{aligned} \tag{3.25}$$

Where the subscript i refers to the ith observation, all the variables X_s and Z_s are exogenous, and $E(u_s | x_1,\ldots,x_M; z_1,\ldots,z_M) = 0$ and $E(\eta_s | x_1,\ldots,x_M; z_1,\ldots,z_M) = 0$. The u_s are assumed to have completely specified continuous marginal distributions and the joint distribution of $(\eta_1,\ldots,\eta_M)$ is also specified. The dependent variable, outcome of interest y_s, is only observed if the sth path is chosen. Thus I_i is a polychotomous variable with values 1 to M and $I = s$ if the sth path is chosen.

$$I = s \text{ iff } I_s^* > \text{Max } I_j^* \tag{3.26}$$

$$\text{Let } \varepsilon_s = \text{Max } I_j^* - \eta_s \qquad \text{where } j = 1,2,\ldots,M, j \neq s$$

It follows that $I = s$ iff $\varepsilon_s < Z_s\gamma$

Lee demonstrates that this approach leads to tractable results if the distribution function of ε_s is specified. If the stochastic components are assumed to be independently and identically distributed with the type I extreme-value (Gumbel) distribution then:

$$\Pr(\varepsilon_s < Z_s\gamma) = \Pr(I = s) = \frac{\exp(Z_s\gamma)}{\sum_j \exp(Z_j\gamma)} \tag{3.27}$$

and hence the distribution function of ε_s is given by:

$$F_s(\varepsilon) = \Pr(\varepsilon_s < \varepsilon) = \frac{\exp(\varepsilon)}{\exp(\varepsilon) + \sum_{j=1, j\neq s}^{M} \exp(Z_j\gamma)} \tag{3.28}$$

For each path s we now have the model:

$$y_s = X_s\beta_s + u_s \qquad \text{iff } \varepsilon_s < Z_s\gamma, \qquad \text{otherwise } 0. \tag{3.29}$$

The distribution function of ε_s specified in equation (3.14) can be transformed to a standard normal random variable with zero mean and unit variance allowing us to estimate the model using the two-step estimate similar to the previous section.

$$\text{Let } \quad \varepsilon_s^* = J_s(\varepsilon_s) \equiv \Phi^{-1}[F_s(\varepsilon)] \tag{3.30}$$

The necessary condition is $\varepsilon_s < Z_s\gamma \Leftrightarrow \varepsilon_s^* < J_s(Z_s\gamma)$ and it is possible, using OLS, to estimate the outcome equation:

$$y_s = X_s\beta_s - \frac{\sigma_s\rho_s\phi\,[J_s(Z_s\gamma)]}{F_s(Z_s\gamma)} + v_s \tag{3.31}$$

After substituting $\hat{\gamma} = \gamma$; and $\sigma_s^2 =$ Var (u_s) and ρ_s is the correlation coefficient between u_s and ε_s^*. The conditional Logit model is used to obtain the preliminary estimate of γ. A number of studies that apply polychotomous-choice models to youth training are reviewed in the next chapter (for example, see Dolton, Makepeace and Treble, 1994b or Payne, 1995).

Duration Models

Duration models can be applied to assess the effect of training on the probability of gaining employment over time, rather than at a specific point in time. This class of models, also known as *survival* or *hazard* models, have as the variable of interest the length of time between 'events' (Goldstein, 1995, p. 125). Let T be the time taken for an individual to obtain employment. Suppose it has a continuous distribution $f(t)$, where t is the realisation of T, the cumulative probability can then be written as:

$$F(t) = \int_0^t f(s)ds = \Pr(T \leq t) \tag{3.32}$$

In practice however, most studies use the *survival function*, which estimates the probability that the spell is at least t, written as:

$$S(t) = 1 - F(t) = \Pr(T \geq t) \tag{3.33}$$

It is then possible to plot the survival functions (labelled Kaplan-Meier plots) for participants and non-participants thus allowing the aggregate features of the data to be compared. They can be adapted to take into account other dynamic features, such as the time spent in training, or type of destination, for example see Dolton, Makepeace and Treble, 1994b, figure 2 where the time to a 'good' job is estimated. The sub-samples used in the Kaplan-Meier plots may not, however, be directly comparable due to underlying differences between participants and non-participants. To take into account observed heterogeneity assume that a spell of unemployment has lasted until time t, it is then possible to determine the probability that it will end in the next short time interval Δ:

$$l(t, \Delta) = \Pr(t \leq T \leq t + \Delta \mid T \geq t) \tag{3.34}$$

We can then derive the *hazard rate* to characterise this aspect of the distribution:

$$h(t) = \lim_{\Delta \to 0} \frac{\Pr(t \le T \le t + \Delta \mid T \ge t)}{\Delta}$$

$$= \lim_{\Delta \to 0} \frac{F(t+\Delta) - F(t)}{\Delta S(t)}$$

$$= {f(t)}/{S(t)} \tag{3.35}$$

The hazard rate *h(t)* represents the instantaneous risk, in effect the conditional probability of someone unemployed at time *t* ending unemployment in the next (small) unit interval of time - in other words the hazard, or risk, of leaving unemployment. The simplest model is one in which the hazard rate does not vary over time, that is $\lambda(t) = \lambda$, characteristic of a process that does not have a memory. In this case the duration time is specified as an exponential distribution, $f(t)\lambda \exp^{-\lambda t}$ $(\mathrm{t} \ge 0)$, which gives $h(t) = \lambda$, so that the hazard rate is constant and the survival function $S(t) = \exp^{-\lambda t}$.[58] In general, however, the hazard rate will change over time and a number of alternate parametric forms have been proposed based on distributions ranging from Weibull to log-logistic (for examples see Greene, 1993, Table 22.6, for a more detailed survey Cox and Oakes, 1985 or Lancaster 1990).

The parametric approaches are relatively simple to estimate but by imposing such restrictive structures on the data the models may distort the estimated hazard rates. Consequently, it is possible to apply fewer restrictions in the form of a non-parametric model such as the Kaplan-Meier survivor function, however a more frequently applied duration model is the Cox Proportional Hazards model, a semi-parametric approach that is used to analyse the effect of covariates on the hazard rate. The proportional hazard model can be defined as

$$h(t;\eta) = \lambda_0(t)\exp^{\eta} \tag{3.36}$$

---

[58] Greene (1993, p. 718) notes that the exponential distribution has been used to model the time to failure of electronic components because of the 'memoryless' property of the distribution.

Where η denotes a linear function of explanatory variables ($X_i\beta$). It is assumed that $\lambda_0(t)$, the baseline hazard function, depends only on time and all the other variation between units is incorporated in the linear predictor η, although the components of η may also depend upon time (Goldstien, 1995, p. 127). In other words, the model does not have a constant because λ_0 is an individual specific constant. Consequently, in principle this parameter should be estimated for each observation, however, Cox's partial likelihood estimator provides a method of estimating the β coefficients without estimating the baseline hazard. Following, Greene, suppose the sample contains K distinct exit times, $T_1,\ldots,T_K$. Then, at any given time T_i, the associated 'risk set' denoted is made up of all individuals whose exit is at least T_i, that is, the set of individuals who have not yet exited just prior to that time. For every individual j remaining in the risk set R_i, $t_j \geq T_i$, and the probability that an individual exits at time T_i, given that exactly one individual exits, is:

$$\Pr[t_j = T_i \mid R_i] = \frac{\exp^{X_i\beta}}{\sum_{j\in R_i}\exp^{X_j\beta}} \tag{3.37}$$

The condition that only one individual exits thus removes the baseline hazard and if this holds for each distinct time T_i, the partial log-likelihood is thus:

$$L = \sum_{i=1}^{K}\left[X_i\beta - \sum_{j\in R_i}\exp^{X_j\beta}\right] \tag{3.38}$$

The Cox model is very flexible because it can readily incorporate time-varying covariates and effectively ignores censoring. This flexibility does, however, come at the cost of estimating various quantities of interest such as the median period of survival and predicted exit time. The next chapter reviews several applications of duration analysis in the training context (for example, see Dolton, Makepeace and Treble, 1994c or Mealli, Pudney, and Thomas, 1996).

3.5.4: Criteria of Interest besides the Mean

The majority of training (and other social) programme evaluations focus on the mean programme impact, for a given sample population under given circumstances. This approach is sometimes labelled 'black box', as it provides no information on the distribution of gains across the population of interest (Ham and LaLonde, 1992; Friedlander, Greenberg and Robins, 1997; Heckman and Smith, 1996). For example, even though the estimated mean programme impact may be zero, it is of interest to know the proportion of participants who benefited from that programme, that is

$$\Pr(y_1 > y_2 | d = 1) = \Pr((y_1 - y_2) > 0 | d = 1) \tag{3.39}$$

Moreover, in a democratic society, it is unlikely that the same weight would be assigned to a programme that benefits only a few people as one with widespread benefits. Indeed a negative mean programme impact may even be tolerated if it can be shown that the majority of participants gain from the programme (Heckman, Smith and Clements, 1997, pp. 490-491). This is more likely to be the case when the benefits from a programme are not transferable, as is the case for training. Similarly, assuming programme costs are borne by taxpayers, it is of interest to know the proportion of the population as a whole that would benefit from the programme, that is:

$$\Pr(y_1 > y_2 | d = 1)\Pr(d = 1) = \Pr(y_1 - y_2 > 0)\Pr(d = 1) \tag{3.40}$$ [59]

With respect to assessing the distributional impact of interventions, Dinardo, Fortin and Lemiuex (1996) recently developed a semi-parametric procedure to analyse the role of institutional and labour market factors on changes in the distribution of wages. Their procedure, which uses an adaptation of the weighted kernel density function, yields a 'visually clear representation' of precisely where in the distribution various factors, such as the minimum wage and de-unionisation, have their greatest

[59] In the broader political economy context, Heckman et al. (1997, p. 491) note that social programmes that result in large but narrowly distributed gains provide an incentive for vested interest groups to actively support the programme in contrast to those in which the gains are more widely distributed. Heckman and Smith (1998) look at the evaluation of the welfare state applying different social welfare functions.

impact. For example, the authors observed that the decline in the real value of the minimum wage in the United States between 1979 and 1988 explained a substantial proportion of the increase in inequality particularly for women and others in the lower tail of the income distribution (ibid. pp. 1014-1019). It is feasible that this technique may be extended to incorporate the impact of training programmes on changes in the distribution of income.

Alternatively, to determine the programme impact on those who benefit the most (and the least) from participation, we can examine different quantiles of the impact distribution. The method of quantile regression, introduced by Koenker and Basset (1978), can be used to characterise the entire conditional distribution of a dependent variable (training impact) given a set of regressors. In other words, it allows the determination of any within-group variation in the estimated impact of training. Following Buchinsky (1997) the standard model can be written as:

$$y_i = X_i\beta_\theta + u_\theta \quad \text{with} \quad \text{Quant}_\theta(y_i|X_i) = X_i\beta_\theta\,, \qquad (i = 1,\ldots,n) \tag{3.41}$$

Where X_i and β_θ are $k \times 1$ vectors, and $X_{i1} = 1$. $\text{Quant}_\theta(y_i|X_i)$ denotes the conditional quantile of y_i, given X_i. β_θ, defined by Koenker and Basset (1978) as the θth regression quantile, is then the solution to the minimisation problem:

$$\min_\beta \frac{1}{n}\left[\sum_{i:y_i \geq X_i\beta}\theta|y_i - X_i\beta| + \sum_{i:y_i \geq X_i\beta}1-\theta|y_i - X_i\beta|\right], \qquad 0 < \theta < 1 \tag{3.42}$$

Other criteria of interest are possible depending on the design and objectives of the programme being assessed. For example, if the programme specifically aimed to improve the earnings of low-income groups, it would be of interest to learn how the distribution of gains depends on the earnings (outcome) in the base, non-participant state. For selected points in the non-participant outcome distribution (y*) this could be evaluated by estimating $F(y_1 - y_2|d = 1, y_2 = y^*)$. Similarly the effects of delivery, advertising and subsidies, or the role of qualifications may impact on the criteria for evaluating programme effects.

3.6: Evaluating Alternative Non-experimental Methods

Non-experimental techniques remain the central tool for the assessment of training schemes in Australia and the UK. One of the main arguments against their application, however, is that there is no uncontroversial way to choose between alternative methods. This view has been formed on the basis of a series of influential evaluation studies undertaken in the US, beginning with LaLonde (1986), that utilise experimental data to assess alternative non-experimental estimators. LaLonde demonstrated that different estimation techniques produced results that were widely divergent and also at variance with the experimental estimates. He claimed that available tests for choosing between the econometric models were ineffective and that social experiments were the only viable evaluation methodology at the time (op cit., p. 614). Proponents of non-experimental methods counter that social experiments are not bias free and, as such, may not be an appropriate benchmark. They also observe that it is not surprising that different models generate different estimates as the models rest on different assumptions, which necessarily effect impact estimates. In fact, only in the absence of systematic differences between participants and non-participants, would alternative non-experimental methods produce the same impact estimate within the relevant sampling variation (Heckman and Hotz, 1989, p. 863).

The lesson we can take from the US experience is the need to adopt some form of model selection strategy that will enable the appropriate choice, of the correctly specified model, for the programme being evaluated. The appropriate model will depend on the available data set, the selection processes underlying the programme of interest and knowledge of the different non-experimental estimators. Our knowledge of the behaviour of these estimators can be enhanced through the evaluation of alternative non-experimental models, here three different approaches are considered. The first two use data sources with known distributional properties, experimental and artificially generated data respectively, to assess the relative performance of non-experimental estimates. While the third approach, sensitivity analysis, applies different non-experimental methods to a common database in order to make any trade-offs explicit.

3.6.1: Comparison with Experimental Estimates

LaLonde (1986), Fraker and Maynard (1987), LaLonde and Maynard (1987), Friedlander and Robins (1995), Heckman and Smith (1999) and others have all utilised data generated from social experiments to indirectly assess the validity of alternate non-experimental estimates of training programme impact. The procedure generally adopted involves the application of each non-experimental method to the experimental data and comparing the results with the experimental estimates. Significant differences between the estimates are taken to mean that the econometric model generating the non-experimental estimates is mis-specified (LaLonde, 1986, p. 611). However, in order to show that non-experimental impact estimates are inefficient one must assume that an impact estimate from a randomised experiment is the true programme impact free from any bias (Friedlander and Robins, 1995).

The approach was pioneered by LaLonde (1986) who set out to assess the validity of the non-experimental estimation techniques used in the evaluation of a temporary employment programme in the USA. He compared the corrected estimates of regression parameters, obtained from non-experimental designs with each other and with those derived from a randomised field experiment, the National Supported Work Demonstration (NSW).[60] Although two-step estimators performed better than one-step techniques, LaLonde concluded that non-experimental techniques, on the whole, failed to replicate the experiment's results and were vulnerable to specification errors.

It is generally agreed that successful experimental evaluations, with random assignment to programme and control groups, are more likely to provide unbiased impact estimates than non-experimental evaluations. In other words, a non-experimental evaluation should aim to obtain an estimate equal to that (asymptotically at least) obtained in a successful or *ideal* experiment (Moffitt, 1991, p. 312). The limitations to *field* experiments, on the other hand, were highlighted in

---

[60] Analysis of the NSW demonstration experiment, also used by Fraker and Maynard (1987) and LaLonde and Maynard (1987), played a pivotal role in the subsequent development of social programme evaluation in the USA, where administrative authorities now favour experimental designs.

section 3.4. Moreover, the LaLonde study has specifically been criticised on the grounds that it tested only a limited range of non-experimental models imposed overly stringent distributional tests and failed to take into account the choice based nature of the sample (Heckman and Smith, 1996, pp. 80-81). Differences between the estimates were described simply as 'large' and not statistically tested (Bell et al., pp. 75-76). The NSW dataset was also restricted in terms of sample size and descriptive variables severely limiting the application of available non-experimental estimators. Significantly it contained insufficient geographical information to place participant and comparison groups in the same labour market, which was identified as a key factor impacting on estimates by Friedlander and Robins (1995), discussed below.

Heckman, Hotz and Dabos (1987), utilising the same dataset as LaLonde, illustrated that some non-experimental methods did result in estimates that closely approximated the experimental estimates (see also Heckman and Hotz, 1989). Heckman et al. found that by invoking an appropriate model selection strategy based on standard specification tests, all but the non-experimental methods that reproduce the inference obtained by experimental methods are eliminated. They identified that a linear regression model, which adjusts for differences in the characteristics of sample members observed just prior to their entry into the programme, produced estimates that closely approximated the experimental estimates in this setting. They point out that different non-experimental estimators will naturally produce different estimates because they make different assumptions about the way people are selected into training and the properties of the error term in the estimating equation (Heckman et al., 1987, p. 417).

Friedlander and Robins (1995) used experimental employment programme data from the Work Incentive (WIN) Demonstration in the USA. They drew data from four different States, which allowed them to focus on how drawing comparison groups from different geographical locations impacts on non-experimental estimates. They also attempted to test if the econometric models successfully accounted for differences between participant and comparison groups except for those introduced by the programme. To do this the authors applied the models to the pre-programme period to see if each model "...'correctly' predicts no differences in outcomes between participant and comparison groups" (p.928). The outcome of interest in the study was employment status and this may however, have impacted on the power of this test, as the fit of the model in the pre-program period is not necessarily related to its validity in the follow-up period. This is because programme eligibility frequently corresponds with deterioration in the circumstances of potential applicants.[61] In the youth context, especially for teenagers, this test is further limited because relatively few members of the sample will have any pre-programme labour market experience with most having only recently left school.

The study found that when comparison groups were drawn from different states the resulting non-experimental estimates were substantially different to the experimental estimates. More crucially the specification test employed by Friedlander and Robins (1995) did not significantly improve the relative position of the non-experimental estimates. When the comparison groups were drawn from the same state as the programme sample the resultant non-experimental estimates were much closer to the experimental estimates. Again the specification test employed did not significantly improve the estimates however it did reject some of the poorer performing estimates. The authors did not apply distributional tests and only tested a limited range of non-experimental estimators concentrating on the linear regression model identified by Heckman and Hotz (1989) as the most efficient estimator in a similar setting. Critically, however, their research found that matching did not significantly improve the estimates. They do not dismiss the use of non-experimental estimators but urge

[61] This is similar to the decline in mean earnings in the period prior to programme entry, observed by Ashenfelter (1978), Ashenfelter and Card (1985) and others, sometimes called 'Ashenfelter's dip'. This severely impacts on commonly used difference-in-difference designs, which inevitably leads to upward bias of training impact in terms of earnings. Heckman and Smith (1999), using JTPA data, found evidence of a pre-programme dip in earnings for all demographic groups included in their analysis of evaluation strategies, although it appeared to be transitory for adult males. Barnow (1987) noted that treatment of the pre-programme earnings played a substantial role in estimates of programme impact. Heckman and Smith (1999) note that estimates are highly sensitive to the timing of pre-programme earning measures (selection of base year). In general, youth datasets in the UK contain few pre-programme labour market observations because the group targeted has been recent school-leavers (16-19 years). An exception is the National Child Development Survey (NCDS), as used by Baker (1990) to assess private sector training in the UK, which allows later entry into training to be evaluated.

caution in their application and encourage further research into the validity of comparison groups.

Heckman and Smith (1999) used data on the NJS (National JTPA Study) control group to assess the validity of two different comparison groups, which may have been applied in non-experimental analyses of training under the JTPA. Examining eligible non-participants (ENPs) and a group of 'eligibles' drawn from the 1986 Full Panel of the Survey of Income and Programme Participation (SIPP), their analysis suggested that conditioning on programme eligibility status did not result in the desired counterfactual relative to the experimental control group. They also found little evidence that the opportunity cost of earnings forgone, suggested by Heckman (1978) and Heckman and Robb (1985), motivated programme participation, rather changes in labour force status tend to predict programme entry. The study also assessed the relative performance of two non-experimental estimators, linear matching and a non-parametric conditional difference in differences estimator introduced by Heckman, Ichimura and Todd (1997). Both procedures reduced, but did not eliminate, the extent of selection bias in impact estimates of training on earnings.

The experimental comparison approach is, however, limited to those countries that allow social experimentation and hence it is not, as yet, an option in either Australia or the UK. The comparison of non-experimental and experimental estimates elsewhere, however, has flagged a number of potentially serious shortcomings that need to be considered in evaluation practice. In particular comparison groups need to be chosen carefully. Friedlander and Robins (1995) suggest that, at a minimum, "...studies must demonstrate the similarity of local conditions as a prerequisite to establishing the validity of the comparison [group]" (p. 935). Heckman and Hotz (1989) propose that specification testing should be standard practice. In general, the studies emphasise the need for greater understanding of the determinants of the participation decision process. All of the studies surveyed, however, used earnings as the only outcome of interest; clearly comparative assessment based on other outcomes of interest would be of value. A range of experimental and non-experimental results from US training programme evaluations are reported and discussed in chapter 4.

3.6.2: Comparisons using Simulated Data

A number of studies have used artificially generated data to assess the relative performance of alternative non-experimental estimators. This approach has considerable flexibility and the testing of a wide range of assumptions is plausible. For example, Paarsch (1984) uses Monte Carlo experiments to compare the performance of different censored regression models including the standard Tobit maximum likelihood estimator (Tobin, 1958), Powell's least absolute deviation estimator (Powell, 1982) and Heckman's two-step method (Heckman, 1979). The Monte Carlo experiment allowed him to address the effects of using different distributional assumptions, small (50) to medium (200) sample sizes and the degree of censoring. His study clearly illustrates the inefficiency of OLS (using only the positive or 'observed' observations, that is $I_i = 1$ using the notation adopted in section 3.5.2) when the data is censored. As might be expected the maximum likelihood estimator (MLE) performed well when the errors are normal, the Heckman and Powell estimators improved with larger sample size. The MLE also performed well when errors are Laplace while the Powell estimator was the most stable when errors are Cauchy.[62] All of the estimators were sensitive to higher degrees of censoring. In a sample selection setting, Nelson (1984) found that the appropriate MLE was again more stable relative to the two-step procedure.

Nawata (1994) adopted a similar framework to that of Paarsch to assess the finite sample properties of the maximum likelihood estimator and Heckman's two-step estimator in the presence of sample selection bias. He utilised the Monte Carlo experiment to examine the effects of the correlation between observables (X_i and Z_i) and unobservables (u_i and ε_i) with 50 percent censoring and a sample size of 200. He finds that Heckman's estimator performs relatively well when the correlation between X_i and Z_i is small, even better than MLE when there is no correlation, but it

---

[62] See the appendix in Lee (1982) for the properties of truncated distributions for a variety of non-normal distributions.

is inefficient in the presence of high correlation. The MLE appears to be more stable in most cases particularly in the unobservables case, when the correlation between u_i and ε_i is large. Nawata recommends that MLE be used when there is a high degree of multicollinearity between X_i and $Z_i\gamma$.

Stolzenberg and Relles (1997) use a large number of simulations drawing on several datasets constructed to approximate different forms of bias. Their analysis finds the Heckman two-step method cannot be relied upon as a matter of course and can in some instances exacerbate the bias. They conclude, however, that the safest approach to problems of sample selection bias is to first understand how non-random selection occurred in the relative data. If, for example, the selection process is consistent with the process described by Heckman (1976, 1979) then that model can be relied upon to provide reasonable estimates. Similarly cautious conclusions can be found in Breen (1996) and Leung and Yu (1996).

The Leung and Yu (1996) study illustrates how artificial data can also be used to enhance our understanding of the operating mechanisms within different estimation methods. They find that the Heckman two-step estimator is effective, in the absence of an instrumental variable, as long as at least one explanatory (X_i) variable displays sufficient variation to induce tail behaviour in the inverse Mills ratio. This is because the Mills ratio becomes non-linear at the extreme values of the permissible values in the single index ($Z_i\gamma$). Hence, given a relatively large range of X_i values it is likely that the data will possess values of the single index, which induce sufficient non-linearity to identify the model.

Although this is a diverse literature, two key findings consistently feature in many of the studies when sample selection is of a known form. First, Heckman's two-step estimator is highly sensitive to size, distribution and the degree of censoring in the sample, indeed it may even exacerbate any bias. Second, in spite of increased computational difficulty, the MLE estimator appears to perform well in a variety of situations and may act as an appropriate benchmark for other estimators, especially when the distribution of the error terms is normal. A caveat appears in Stolzenberg and Relles (1997), however, who observe that there "...appears to be no automatic way to diagnose and correct sample selection bias" (p. 504).

The results from using artificially generated data to evaluate non-experimental methods have been mixed and the literature generally adopts a neutral stance suggesting that caution should be exercised in their application to programme evaluation. The versatility of this approach has allowed non-experimental estimators to be tested in a wide variety of hypothetical scenarios, which has greatly enhanced our knowledge of the behaviour of those estimators. The need to carry out appropriate specification testing is again highlighted, in this case to enhance our knowledge of the form of sample selection inherent in our data.

3.6.3: Sensitivity Analysis

Clearly, the parameter estimates and assumptions of any economic model are subject to change and error. It is important to try to check that our assumptions hold, to transform data to fit the assumptions where possible, and to design and undertake research so as to minimise such problems (Breen, 1996, p. 13). Sensitivity analysis provides a means to investigate these potential changes and errors and their impacts on any conclusions to be drawn from the model. Sensitivity analysis involves the application of different model specifications to a common database in order to reveal more about the construction and interpretation of the resulting estimates as well as aid model selection. The versatility of sensitivity analysis can be gauged by a brief review of some earlier applications.

Mroz (1987) used the Panel Study of Income Dynamics (PSID) to complete a detailed sensitivity analysis examining the impact of different statistical and economic assumptions on the participation of married women in the labour force. Not unlike impact estimates in the training evaluation literature, reported estimates of the response in female labour supply to variation in wage rates and income varied greatly. After constructing a baseline estimation using OLS, Mroz examined the sensitivity of this estimate to different exogeneity assumptions, to different methods of statistical control for self-selection into the labour force, and the impact of controlling for taxes. He found that controlling for self-selection into the labour force substantially reduced upward bias in estimates produced by several different exogeneity assumptions. That is, unobserved omitted variables impact on the labour market participation and hours of work decision. Additionally, Berndt (1992) notes the considerable interest of Mroz's study because it backed a challenge to the conventional wisdom of labour economists that the responsiveness of female labour supply to changes in wage rates and income is larger than that of men (pp. 637-638).[63]

Dickinson, Johnson and West (1987), employed sensitivity analysis in an attempt to reconcile the wide range of reported non-experimental impact estimates of training programmes administered under the Comprehensive Employment and Training Act (CETA) in the USA. The authors noted, "Because of the numerous methodological choices made in the various studies, it is difficult to identify which specific factors are responsible for the different impact estimates. It is important, therefore, to examine empirically which methodological factors affect the net impact estimates" (p. 453). They examined the sensitivity of their initial results to, the matching technique used, the choice of sample frame used to draw the comparison groups, the alignment of the treatment and control samples, and the estimation model used. Impacts proved quite robust to different matching procedures but, in contrast, were relatively sensitive to sample construction choices. For example, estimates proved particularly sensitive as to whether individuals without recent labour market experience were included in the comparison group or not. Alignment factors were also sensitive in terms of the accuracy of pre-programme earnings. With respect to model choice, Dickinson et al. only tested the one difference in differences estimator choosing to concentrate on the timing of pre-programme earnings when specifying the participation decision. The estimates for women were more sensitive to the different specifications than those for men.

In general, the Dickinson et al. (1987) study showed that the making of seemingly minor choices in the construction of the estimates resulted in widely divergent estimates of the impact of the CETA programmes. Riddell (1991, p. 60) observes that a contributing factor to the diversity in estimates is however, the fact that

---

[63] This confirmed the stand taken in a series of articles by Nakamura and Nakamura, see for example Nakamura and Nakamura (1985), "Dynamic Models of the Labor Force Behaviour of Married Women Which Can Be Estimated Using Limited Amounts of Past Information," *Journal of Econometrics,* 27(3), 273-298.

earnings in the comparison group were measured on a calendar year basis whereas the earnings of participants were measured on a fiscal year basis (see also Warburton, 1996, p. S107). Heckman and Robb (1986, p. 67) also caution that part of this variability in estimates may be due to the imposition of extra conditions not required to identify the parameters of interest.

In response to an 80 percent cut in the youth component of the JTPA (successor to CETA) budget following unfavourable experimental impact estimates, Heckman and Smith (1997) assessed the sensitivity of experimental estimates of youth training under the JTPA to a variety of modelling choices. These include the study of different training centres, techniques for combining estimates from different training centres, the treatment of outliers, the construction of earnings measures, the treatment of attrition and, how control members that substitute into alternative training schemes are dealt with. Again, seemingly minor choices in the construction of the data set and method applied resulted in widely divergent estimates of programme impact. Heckman and Smith (1997, p. 32) report that impact estimates on the earnings of disadvantaged male youth were extremely fragile and that it is more likely that JTPA has a zero impact for this group rather than a negative one. Their study thus indicates that experimental impact estimates, like their non-experimental counterparts, require careful interpretation if they are to provide a reliable guide to policy makers.

The latter two studies illustrate that divergent estimates in training evaluation, for the most part, can be explained by the different methodological procedures used. This prompted Dickinson et al. (1987) to emphasise "… the important role of sensitivity analysis as an integral part of any net impact analysis" (p. 465). Sensitivity analysis cannot solve the problem of sample selection bias but, in combination with established testing procedures, it can aid in selecting the appropriate estimator for the relevant sample. This also rests on increased understanding of the behaviour of different estimators in various settings and it is in this area that we can draw on the results from comparisons using simulations and social experiments. Sensitivity analysis is now widely used in the applied economics literature, at least in some limited form. Leamer (1983, 1989) was instrumental in raising the profession's awareness of this approach but, to a great extent, it is now implicitly a part of good practice. For example, Winship and Mare (1992, p. 347) suggest that because evaluation results may depend on the method used, it is important to make explicit the assumptions underlying the model and present estimates using a variety of methods.

3.7: Chapter Summary

This chapter has identified the necessary ingredients for efficient impact analysis in the field of training programme evaluation. The principle challenge facing the analyst is to construct a realistic counterfactual as failure to do so renders estimates unreliable, due to sample selection bias. Experimental methods offer the most promising solution to the evaluation problem but the implementation of a successful social experiment faces a number of significant limitations. These range from the practical difficulties arising from randomisation and substitution bias through to the ethical conflict of having to deny eligible applicants places on programmes. At the current time, social experiments are not a valid option for the evaluation of training programmes in either of our countries of interest or many other modern economies.

The fundamental finding of theoretical investigation into non-experimental methods for training programme evaluation is that there is a high likelihood that OLS estimates will be biased because of sample selection. Consequently, a wide variety of alternative models have been proposed to correct for this bias, however it is clear from this chapter, that no single method will always solve the evaluation problem. For each of the models presented, at least one minimal identifying assumption always needs to be made. As shown by Heckman and Smith (1995) this, in fact, holds for all non-experimental methods and because minimal identifying assumptions cannot be tested with data, model choice can be complex and controversial.

A number of factors need to be considered when choosing an appropriate non-experimental model including, the nature of the training programme, objectives of the programme, the availability of appropriate data, and the institutional setting. Crucially, it is also important to utilise and build on our knowledge of how different estimators behave in different settings. For example, comparisons in the performance of non-experimental estimators with results derived from experimental and artificially generated data emphasise a key role for specification and distributional testing. Sensitivity analysis, on the other hand, allows checks for consistency of different assumptions to be made within the context of the programme being evaluated. Drawing on the studies that have made comparisons using simulations and social experiments to help verify the validity of different estimators can strengthen this and is thus the methodology applied in the empirical section of this thesis (chapter 6).

The next chapter presents an overview of recent institutional developments associated with the provision of youth training in Australia and the UK. It also reviews the relevant empirical literature on youth training programme evaluations by various outcomes of interest and presents an overview of relevant datasets. The main focus is on the evaluations made of government-supported programmes in the UK, which will illustrate the application of many of the models detailed in this chapter. Moreover, the reported estimates of programme impact in the UK clearly illustrate the wide range in estimates inherent in non-experimental evaluations of similar programmes using seemingly similar datasets. The non-experimental UK studies are also compared and contrasted with a selection of programme evaluation results from the USA, where experimental studies are permitted.

Chapter 4: Institutional Developments in the Youth Labour Market and Survey of the Empirical Literature

4.1: Introduction

> "Show me a skilled individual, a skilled company, or a skilled country and I will show you an individual, a company or a country that has a chance to be successful. Show me an unskilled individual, company or country and I will show you a failure in the 21st century" (Thurow, 1994, p. 52).

Government intervention in the youth labour market to promote training has seemingly become a permanent feature of most industrial economies. The expansion of education and training for young people is one of the key platforms in modern government policy underpinning the aim of developing internationally competitive industry. It is also motivated by the widely held perception that firms under-invest in training because trained workers might leave, or even be 'poached' by other firms. Historically, vocational education and training systems, both in the United Kingdom and Australia, have been viewed as inadequate relative to those of their competitors. Arguably, they lacked the necessary flexibility to meet the demands of substantial structural and technological change that characterised industrial economies in the second half of the twentieth century. This resulted in significant disequilibrium in the youth labour market, which was exacerbated by the 1970s recession. The response was an increase in the level of government intervention with the aim of addressing both youth unemployment and perceived skill deficiencies. Specifically, in the United Kingdom government training policy has, in the last three decades, moved away from simply regulating employers to the subsidisation of large-scale youth training programmes. This has resulted in considerable reshaping of the youth labour market including an increase in the variety of choices young people face in the school-to-work transition. In contrast, until recently, Australia continued to rely on regulatory measures and employer incentives to provide adequate training, chiefly through the apprenticeship system. This has been supported by institutional arrangements to assist in the delivery of training and a plethora of small-scale youth training programmes targeting the long-term unemployed and disadvantaged groups.

Continued government commitment together with changes in the design and delivery of training programmes increases the pressure for more accurate evaluation of such programmes.[64] This chapter reviews existing economic evaluations of different government initiatives designed to promote youth training in the United Kingdom (UK) and Australia. The emphasis is on youth training programmes in the UK, where existing evaluations clearly illustrate the variability of estimated programme effect that can arise from the application of non-experimental methods even when applied to broadly similar programmes using seemingly similar datasets.[65] This variation in estimates underlies much of the criticism levelled at non-experimental evaluation methods, however, a number of other factors also contribute to this variation including different institutional settings, changing economic conditions, variation in delivery across programmes and regions, and the application of different evaluation methodologies over time. This chapter thus also presents an overview of these factors in the UK and Australia focusing on the main economic policy issues and institutional developments in youth training over the last 30 years. One of the key developments within this period, common to most industrial countries, is the significant convergence between academic and vocational education, which makes it increasingly difficult for the evaluator to determine what actually constitutes training (Stern and Wagner, 1999). The increased numbers of young people participating in school-based vocational training options has accentuated this blurring of the boundary between education and training.[66] The frequent absence of any formal structure and consequential variability in the delivery, accreditation and setting of training standards also hamper definition. Moreover, in terms of outcomes, it is often difficult to disentangle the impact of training from other influences, exacerbated by extensive overlap in public and private sector provision.

---

[64] Evidence of the UK Government's ongoing commitment to investment in training and education and proposed monitoring of performance and delivery standards can be found in the recent White Paper, (1999) *Learning to Succeed: A New Framework for Post-16 Learning*. Australia's aims and initiatives for vocational education and training are contained in the Australian National Training Authority publication, ANTA (1999) *A Bridge to the Future: Australia's National Strategy for Vocational Education and Training.*

[65] There are a number of existing reviews of government youth training provision and programme evaluations in the UK, see for example Dolton (1993), Bradley (1995) or Deakin (1996).

[66] Favennec-Hery (1996) goes further and argues that the boundary between training and work is also blurred and fraught with definitional difficulties "...'work and training' are best approached not as static objects of study, but rather as an ongoing process" (p. 666).

Early evaluation attempts tended to rely on aggregate data relating to the relative employment probabilities and mean post-programme earnings of participants and non-participants. This approach does not, however, allow assessment of the effect of training in terms of individual earnings or employment prospects at the same time as controlling for interpersonal differences. Nor does it adequately deal with problems arising from sample selection (Dolton 1993). Aggregate measures do however continue to be widely relied upon by most government agencies for monitoring purposes but, as established in chapter 3, effective evaluation requires the estimation of the counterfactual for programme participation. The advent of large-scale individual cross-section and longitudinal data in the 1970s allowed the application of more sophisticated econometric techniques designed for this purpose. These techniques, sometimes labelled second-generation studies, attempt to control for differences between participants and non-participants as outlined in chapter 3. Focusing on microeconomic evaluations, which estimate the impact of the programme on the individual participant, the studies reviewed in this chapter illustrate most of the models presented in chapter 3 in terms of the impact on earnings and employment prospects. A number of alternative approaches are also briefly discussed.

The remainder of the chapter is structured as follows. Section 4.2 details the development of recent government policy initiatives on youth training in the UK. It highlights the changing objectives and design features characteristic of the key programmes introduced in response to the changing state of the labour market. These include the introduction of the Youth Opportunities Programme (YOP) in 1978, which represented a radical departure in national training policy setting in place many features that remain part of current programmes. Section 4.2 also details the provision of youth training as it has subsequently evolved through different institutional permutations, each introducing distinct objectives, and reflecting the economic and political circumstances at the time of their inception. However, the key aim remains to enhance youth employability. Significantly, as the programmes have developed government agencies have taken increasing responsibility for the training content, which has moved toward greater standardisation in delivery, reflected in the adoption of a national qualification framework. The focus of this section is thus on the main economic policy issues and institutional developments in youth training provision over the last 30 years. The next section focuses on attempts to evaluate the impact of such intervention in the youth labour market.

The largest and most ambitious of the UK training programmes was the Youth Training Scheme (YTS), which has also been the most intensively evaluated of the different programme permutations. Section 4.3 discusses such programme evaluations in more detail beginning with a brief overview of the main datasets employed in the studies in section 4.3.1. Section 4.3.2 then reviews reported estimates of the impact of UK training programmes in terms of subsequent earnings, which vary between minus 16 percent and plus 4 percent, dependent on the type of individual. To allow some comparison with experimental estimates of earnings impact, a sample of studies from the USA are also presented to illustrate that the spread in estimates is not entirely mitigated when experimental data is available. Section 4.3.3 then presents the reported estimates of the impact of training on the subsequent probability of gaining employment, which range from slightly negative to an improved probability of 21 percent. Estimates of post-training earnings impact and employment prospects have dominated the evaluation literature however, section 4.3.4 surveys a selection of evaluations that have examined 'qualifications gained' as an alternative outcome of interest. Section 4.4 then presents a summary of the different UK studies on the basis of model choice and specification to highlight the possible impact of decisions made by the researcher on resulting impact estimates.

Recent developments in the Australian youth labour market, where the government has traditionally adopted a less interventionist stance, are detailed in section 4.5. It is shown that the Australian training system shares many features in common with the UK, reflecting its British heritage. This is particularly apparent in the traditional apprenticeship system. However, while the UK has increasingly moved toward subsidised training schemes in recent decades, Australia has attempted to modify its apprenticeship system through institutional and regulatory reform. At the same time the Australian reforms have embraced many of the same objectives as British reform, including greater occupational coverage, increased standardisation of delivery, and a greater range of formal vocational qualifications. Australia has also utilised a range of small-scale youth training programmes, similar to the British interventions, with the aim to assist disadvantaged labour market groups, particularly Indigenous Australians. Significantly, in the last decade greater co-operation across States and Territories has allowed the Commonwealth (Federal) Government to successfully introduce reforms in the governance of its vocational education and training institutions and develop a nationally consistent qualification framework. This has led to increased intervention although, as detailed in section 4.5, much of this revolves around a more flexible apprenticeship scheme, the so-called *New Apprenticeship* system.

Section 4.6 reviews a selection of recent evaluations of the different Australian youth training initiatives although only a few of these have adopted the rigorous econometric approach common to recent UK evaluations. Several studies specifically assess the impact of training on the labour market prospects of Indigenous Australians reflecting the relevant policy emphasis. However, given the reliance on employer-led initiatives and small-scale programmes, it is not surprising that most of the Australian evaluation literature focuses on private sector training. This is also reflected in limited data availability on government assisted training placements, which tends to favour aggregate, rather than econometric, assessments. This section (sub-section 4.6.1) thus also includes an overview of relevant Australian data sources that might be used for evaluation of labour market interventions in future research especially as the New Apprenticeships become more established and readily identifiable in those datasets. Finally, a chapter summary is provided in section 4.7.

4.2: Training Provision in the United Kingdom

Training in the UK traditionally centred on 5 to 7 year apprenticeships served by the individual with a single employer. Apprenticeships were almost exclusively based in the manufacturing sector and funded by the employer but had little uniformity in standards or qualifications, and were predominately the preserve of male youth. The sectoral Industrial Training Board (ITB) system, set up in 1964 was charged with improving the quantity and quality of training within industry. The ITB system was based on a grant/levy mechanism without major government funding, working on the principle of rewarding companies who were training properly and penalising those who were not.[67] A pronounced increase in unemployment rates in the 1970s and a shift in the structure of demand for skilled workers in the face of new technology supported criticism that the ITB system was failing to deliver the training that was needed (Deakin, 1996, p. 70). The system was subsequently overhauled by the Employment Training Act 1973, which set up a centralised national training body, The Manpower Services Commission (MSC). The Commission was charged with bringing together vested interest groups (employers, ITBs, unions, etc.), in order to create a coherent national system of training with an integrated approach to labour supply (Dolton, 1993, p. 1263-1264). The MSC implemented a number of small-scale initiatives for unemployed youth, such as the Youth Employment Subsidy and the Work Experience Programme. Such initiatives were counter-cyclical measures with the aim of maintaining employment directly rather than through the provision of training.[68]

In the early 1970s the school-to-work transition for most young people was relatively straightforward with well over half of the 16-year-old cohort entering directly into employment and any training was typically confined to (traditional) apprenticeships. The onset of recession later in the same decade witnessed a dramatic rise in the level of youth unemployment, from less than 10 percent in 1978 to more than 25 percent by 1983.[69] A significant contributory factor was the relative decline of the British manufacturing sector, which significantly curtailed opportunities for traditional apprenticeships. Apprenticeship placements in manufacturing fell from 240,000 in 1964 to 155,000 in 1979 and by 1986 numbered just 63,700 (Dolton, 1993, p. 1261; Blanchflower and Lynch, 1995, p. 235). Moreover, relatively large school-leaver cohorts, reflecting higher birth rates in the 1960s exacerbated the problem. In other words, the youth labour market was subjected to a substantial increase in the supply

---

[67] One of the original ITBs, the Engineering Construction ITB, still survives with responsibility for determining training needs within its sector and it retains the statutory power to raise a training levy of up to 1.5 percent of each member firm's payroll (Stevens, 1999).

[68] For more detailed discussion on these earlier initiatives see Deakin, 1996, chapter 5.

[69] Figures 2.1a and 2.1b (chapter 2 of this thesis) plot the youth unemployment rates for the UK between 1978 and 2000.

of labour in the mid 1970s, due to changing demographics, at the same time as its demand weakened (Deakin, 1996, p. 73).

In 1978 responding to what, at the time, was perceived as a temporary excess supply of youth labour, the government introduced the *Youth Opportunities Programme* (YOP). The aim of the programme was to significantly reduce the cost of trainees to employers and to ensure young people gained work experience making them more employable. Participant allowances were set at a rate slightly higher than benefit levels but not so high as to discourage young people from continuing education. In other words, the programme was designed to cause minimum interference to the workings of the labour market. The expectation was that the YOP would be a relatively small-scale programme that would only be required until the economy picked up. However, the youth labour market virtually collapsed in the early 1980s so that numbers entering the programme leapt from approximately 40,000 in 1978 to 368,900 in 1981 (Bradley, 1995, p. 35). The election of the Conservative Government in 1979 led to the dismantling of the ITB system however, in the face of a continued deterioration in the youth labour market, the YOP was continued and by 1981 the programme was absorbing approximately one in three school leavers.

The YOP was not a homogeneous scheme; rather it was made up of different types of provision for different client groups (Bradley, 1995, p. 35). It was initially divided into *work experience placements*, attracting the large majority of participants operating through private sector employers, and *work preparation places*, mainly comprising short (13 week) training courses such as training workshops and community service, many of them operated through the voluntary sector. In 1982 the programme was extended to explicitly include *new training places*, designed to last 12 months with time divided between a sponsor's premises and colleges of further education (Deakin, 1996, pp. 77-79). The introduction of the new training places indicated a shift in national youth training policy, first signalled with the publication of the 'National Training Initiative' by the MSC in 1981. Significantly, this initiative emphasised training as a means of promoting youth employment. The introduction and development of the YOP thus established the principle that all those people aged between 16 and 18 who left school and were not in full-time education or were unable to get a job, should have the opportunity of training (Dolton, Makepeace and Treble, 1994a, p. 263; Deakin, 1996, p. 85). Another important feature of the YOP was the role of the Careers Service, which was assigned the role of 'gatekeeper' to the programme. The Careers Service informed young people in school of the various YOP options and attempted to systematically match any individuals who subsequently became eligible for relevant options based on observed characteristics (Bradley, 1995, p. 36).

As the size of the programme increased a number of weaknesses became apparent. It was widely denounced as a source of cheap labour and post-programme employment placement rates were beginning to decline (Bradley, 1995, p.37). In 1983 the YOP was subsequently replaced with the *Youth Training Scheme* (YTS) firstly as a one year and then (in 1986/87) as a two-year programme. The YTS was designed to draw on the lessons and strengths of the YOP. In particular, its introduction witnessed greater emphasis on the quality of training and the scheme became widespread in many sectors of British industry and commerce covering an increased range of jobs and to a great extent supplanting traditional apprenticeships.[70] Young people did not have to be unemployed to participate in the programme further emphasising that it was designed to assist the school-to-work transition rather than youth unemployment *per se*. Work experience continued to be a major component of the programme but participants were encouraged to gain experience with more than one employer. The training component comprised three months compulsory off-the-job training at a local college of further education or suitable training workshop. The emphasis was on the creation of transferable or general skills in the hope of creating a more highly skilled workforce. The desire to increase opportunities for vocational training, leading to recognised qualifications, motivated the shift to a two-year programme. The MSC's objective was to provide high quality

[70] In a series of articles Gospel documents the changing fortunes of the traditional apprenticeship route, for example see Gospel (1994), which looks at the relative decline of traditional apprenticeships in the UK, Australia and the USA. For more general discussion on changes to the apprenticeship system see Junanker (1987) or Blanchflower and Lynch (1995). Dolton (1993, pp. 1268-1269) notes that the interpretation of apprenticeship figures is complicated following the introduction of YOP and YTS due to possible double counting and increased availability in non-manufacturing sectors.

training, historically confined to a narrow range of apprenticeship occupations, across the entire entry-level occupational spectrum (Bradley, 1995, p. 39; see also Begg, Blake and Deakin, 1991).

Like its predecessor, the YTS continued to offer a variety of different options for different client groups and the Careers Service continued its role placing over two thirds of the entrants onto the programme (Bradley, 1995, pp. 37-38). Approximately 75 percent of entrants were on employer-led training schemes (Mode A) organised by industrial and commercial organisations in both the public and private sectors. The remaining entrants were on schemes run largely by voluntary or non-profit organisations (Mode B). Many of the latter schemes catered specifically for young people disadvantaged either because of physical or learning disabilities and were fully funded by the MSC. A small portion of the Mode B places were in MSC operated Information Technology Centres, which offered training in IT related occupations for which demand was increasing rapidly. The distinction between Mode A and Mode B schemes formally disappeared with the extension of the programme to two years. Young people with special needs were subsequently provided for by means of premium places, which attracted a higher grant to employers (Jones, 1988, p. 56). The task of determining which funding category a young person fell into continued to rest on the Careers Service but scheme administration was devolved to a diverse network of Managing Agents, including a wide range of public and private sector bodies (ibid. p. 56). In addition, employers were encouraged to use YTS to fund the first stage of apprenticeship training and utilise spare training capacity (Bradley, 1995, p. 38). For every two long–term trainees recruited, such as apprentices, employers could recruit three additional trainees funded by the YTS. Bradley argues that this made it possible for employers to use YTS as a screening device for borderline apprenticeship candidates (ibid. p. 38).

At the same time, the Conservative administration adopted employment policy designed to keep the cost of youth labour at an attractive level for employers. This was the aim of the *Young Workers Scheme* (YWS) introduced in 1982, which provided a wage subsidy to employers who recruited individuals aged less than 18 in their first year of employment, and agreed to pay up to a wage ceiling. A wage of up to £40 per week attracted a £15 subsidy and a wage between £40 and £45 a subsidy of £7.50. Approximately 290,000 young employees were covered by the subsidies between 1983 and 1988 (Bradley, 1995, p. 39). Significantly, from an evaluation perspective, many of the YWS beneficiaries were previously on YTS programmes, which is likely to have suppressed post-programme earnings (Hutchinson and Church, 1989). In 1986 the YWS was replaced with the New Workers Scheme (NWS), which was tailored to fit with the extension of YTS to a 2-year programme. Under this scheme employers could claim a £15 subsidy for individuals under the age of 20 employed at a wage ceiling of £55 and increasing to £65 for 20 year olds (Deakin, 1996, p. 117). Implicit in these policy initiatives is the assumption that the wage expectations of young job seekers are unrealistic and need to be ameliorated. To some extent, the low levels of the training allowance in the YOP and YTS programmes can be viewed in the same light (Main, 1987).

In the early 1980s the MSC, employers and other interested bodies became increasingly aware that the problem of youth unemployment was related to both education and training (Raffe, 1990, p. 59). Greater consideration was hence given to the education curriculum with the launch, in 1982, of the Technical and Vocational Education Initiative (TVEI), which was designed to stimulate technical and vocational education amongst 14 to 18 year olds in school. These courses ran in parallel with standard academic courses and could include work experience with the general aim of smoothing the school-to-work transition. The TVEI was extended to all secondary schools in 1987 (see Deakin, 1996, chapter 7). In England and Wales, the National Council for Vocational Qualifications was also established at this time in an attempt to standardise qualifications in business and industry. The Council was charged with establishing the framework for a new National Vocational Qualification (NVQ) system to encompass the many and varied vocational qualifications awarded by more than 300 different bodies in the UK (Deakin, 1996, p. 95-96; see also Robinson, 1996, pp. 78-79). NVQs are awarded at a number of different levels, each level being defined as "a measure of competence of an individual's capabilities to carry out a range of work to standards ... agreed by industry" (Steedman and Hawkins, 1994, pp. 95-96). A higher level of integration between vocational and academic education was achieved following the introduction of the school-based General National Vocational Qualification (GNVQ) system in 1991. GNVQs were designed to be the equivalent of 'A-Levels', which are required for entrance to university. Although popular with young people, enrolment data indicate that GNVQ courses attract students with considerably poorer results in examinations at 15 or 16 years of age compared with A-level students (Payne, 1999, p. 484). Nevertheless, combined with changes to the system of academic assessment, these innovations have resulted in the spread of certification to encompass virtually the whole of the school population and may have contributed to higher staying-on rates in education (Heath and Cheung, 1999; Robinson, 1996).[71]

The convergence between academic and vocational education appears to have proceeded even further in Scotland where a modular system of vocational qualifications leading to a National Certificate was developed following the 1983 Scottish Action Plan. The flexibility of the Scottish modules allowed the mixing of vocational and academic education, while meeting specific employer and student needs, and encouraged greater participation in further and higher education (Stern, Bailey and Merrit, 1996, p. 18). The subsequent introduction of Scottish Vocational Qualifications (SVQs) in 1989 and General Scottish Vocational Qualifications (GSVQs) in 1991 built on this framework using the same accrediting body, the Scottish Vocational Education Council, which works closely with employers and Further Education Colleges. The take-up of both NVQs and SVQs has, however, been relatively slow and predominately confined to government programmes, which require them as a condition of funding to training providers (Raffe, Biggart, Fairgrieve, Howieson, Rodger and Burnisten, 1999). This suggests the main hurdle still confronting vocational education is its low status relative to academic education (Payne, 1999, p. 497).

---

[71] The key reform in the UK's system of academic assessment was the introduction of the General Certificate in Secondary Education (GCSE) in 1987/88 with the aim of raising pupil performance. A concise outline of this and other recent educational reforms can be found in Heath and Cheung (1999). Education staying-on rates increased rapidly in the UK throughout the 1980s and these were detailed in chapter 1 (see figure 1.3a).

In 1989 *Youth Training* (YT) succeeded YTS as the primary vehicle for the delivery of training for young people. YT was modelled closely on YTS with some modifications, the main one being the guarantee of a place for all those aged 16 to 18 years that were without a job (100 percent coverage). At the same time training courses were given more flexibility allowing courses of variable length, with no fixed design framework reflected in different lengths of off-the-job training for different occupations, and local variations in funding arrangements (Bradley, 1995, p. 41; Deakin 1996, pp. 149-150). To facilitate these changes and to encourage greater involvement of employers on a local level, Training Enterprise Councils (TECs) were set up in England and Wales, and Local Enterprise Companies (LECs) were set up in Scotland. The TEC/LECs, which were designed to be run by private enterprise employers, had the aim of tailoring training to local demands.[72] YT witnessed an increased emphasis on NVQ/SVQs with participants encouraged to acquire a minimum of NVQ/SVQ level 2 (broad-based skills: semiskilled occupations). Moreover, funding arrangements also changed with payments being made to firms dependent on the trainees working towards a recognised qualification with a proportion, generally 40 percent, of funding withheld unless the qualification is achieved (Steedman and Hawkins, 1994, p. 95; Bradley, 1995, p. 41). This was consistent with the approach adopted by the Conservative Government to foster the introduction or strengthening of market forces in many areas of state provision.[73] It has, however, led to charges that training provision has become more concerned with outcomes rather than input or processes (Steedman and Hawkins, 1994, p. 100).

---

[72] TEC/LEC funding predominately comes from the administration of government-funded training schemes, typically subcontracting the delivery of this training to a range of organisations in the public and private sectors (Payne, 1999, p. 481). A number of TEC/LECs have piloted *youth credit* schemes whereby young people purchase training under YT thus giving them greater control over training choices, however the scheme did not gain wide support. Note that LECs are also responsible for economic development and come under the umbrella of two central bodies, Scottish Enterprise and Highlands and Islands Enterprise.

[73] Davies and Adnett (1999) investigate the impact of quasi-market reforms on vocational schooling in England and Wales and concluded that they have exacerbated problems generated through inequalities (cultural and social) and has reinforced the cycle of low attaining pupils from non-professional backgrounds opting for vocational routes (pp. 9-10).

Modern Apprenticeships, first available in 1994, increased the emphasis on vocational qualification still further. These extend government assisted training by two or three years beyond what was previously available and lead to a NVQ/SVQ level 3 (key skills: technician, craft and supervisory) or above.[74] At the end of August 1988 a total of more than 218,000 young people had started on modern apprenticeships since the schemes inception. Significantly, modern apprenticeships are available over a wide range of industry sectors, including areas where there has been no tradition of apprenticeship such as retail, accounting and estate agency (see Deakin, 1996; D*f*EE, *News Update*, 574/98).

The latest development in a complex ever-changing system of intervention, *National Traineeships*, were introduced in September 1997, and will eventually supplant YT as the primary vehicle of delivery of youth training in England and Wales.[75] The Traineeships share many of the design features of Modern Apprenticeships but are shorter in duration and focus on the NVQ level 2. In Scotland, YT has been replaced by *Skillseekers*, which was phased in across the country from 1992 and achieved national coverage in April 1996. Scottish Enterprise and Highlands and Islands Enterprise, the two economic development agencies in Scotland, through their networks of LECs administer Skillseekers. It is aimed primarily at 16 and 17 year olds, the youth training guarantee group, who are looking for a work-based learning alternative to post-compulsory education or employment. Training opportunities are also available for those individuals aged between 18 and 24 years, on a discretionary basis, with LECs expected to match provision with local economic needs. Participants on placement with employers receive a training allowance, which is

---

[74] Modern Apprenticeships are available to 16 and 17-year old school leavers, and individuals aged 18 and over (at the TEC/LECs' discretion) that can complete the apprenticeship before the age of 25. One of the criteria for employers is that the apprentice should be given employed status, and thus paid wages, as soon as possible (D*f*EE, *Modern Apprenticeships Factsheet*).

[75] YT continues to exist, appearing in official statistics as 'Other Training', with 148,400 starts in 1988 although numbers have subsequently fallen as policy shifts towards Modern Apprenticeships and National Traineeships/Skillseekers, see also figure 4.1. The emphasis on qualifications gained as an outcome inherent in these more recent programmes is sometimes referred to as 'competency based training' (Stern et al., 1996, pp. 16-22). A more detailed description of the qualifications framework and requisite standards can be found in Raffe et al., 1999, pp. 6-14. For discussion on competency based training in the Australian context see section 4.5 below.

normally £45 to £50 per week. Skillseekers entrants follow an agreed training plan leading to SVQ level 2 or 3 and increasingly have employed status while receiving training.[76] These initiatives have been accompanied by an enhanced role for the Careers Service in ensuring young people take up their entitlement to continued learning after the age of 16 (D/EE, 1999). Higher staying on rates at secondary school and lower unemployment in the 1990s has, to some extent, marginalized youth training programmes as an option in the school-to-work transition. The proportion of the school leaver cohort opting for government supported training remains relatively stable at approximately 10 percent since the early 1990s but the various schemes continue to cater for significant numbers of young people. The combined number of participants in YT, Modern Apprenticeships and National Traineeships in England between 1992 and 1999 are illustrated in figure 3.1. While these programmes will continue to be subjected to changing conditions of market demand and supply, and will naturally evolve in response to technological advances and changing administrative and political considerations, they remain an important component of employment policy.

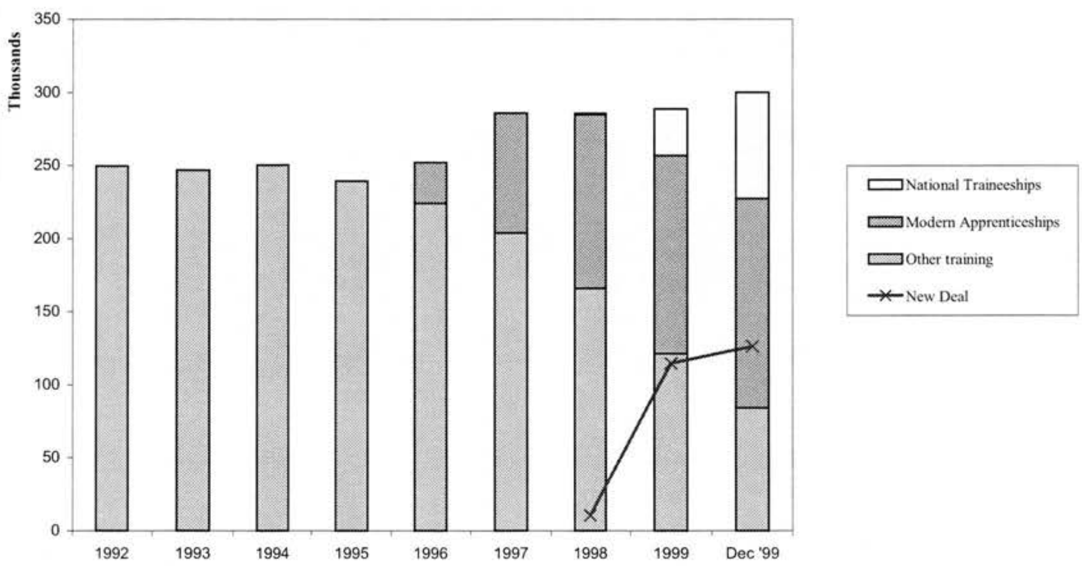
In a parallel development, the so-called *New Deal*, introduced by the newly elected Labour Government in 1997, targets young people who have been unemployed for 6 months or longer. With relatively low starts in the National Traineeship and Skillseekers options, the New Deal has, to a great extent, become the Government's primary vehicle for addressing youth labour market issues. The relatively rapid take-up rate of New Deal places is shown in the line-graph superimposed on figure 4.1. However, given the workings of this initiative, it shares more in common with the YOP than the more recent programmes. Like the YOP the emphasis is on work experience, the tenure is reasonably brief, and training is not well defined. The programme offers four options, all of which include a training element, subsidised employment, returning to full-time education or training, work in the voluntary

---

[76] In March 1999 the Scottish Executive (1999, p. 55) reported that there were 37,600 participants in Skillseekers, including 10,400 on Modern Apprenticeships. Of these 72 percent received a wage from their employer. The Executive has set itself a target of achieving 20,000 modern apprenticeship placements in 2003.

sector, or a place with the Environmental Taskforce. The scheme allows for an initial 4-month 'gateway' for matching candidates to the most beneficial option that may involve a 6-month (or more) subsidy or allowance being allocated dependent upon the option chosen. The subsidised employment element can involve up to £60 per week subsidy in addition to a £750 contribution towards training costs. In return employers are expected to arrange training and release the trainee for the equivalent of one day per week. Significantly, under the New Deal, youth training policy has been expanded to incorporate those individuals aged up to 24-years.

Figure 4.1: Participation in youth training, England: 1992-99



Source: DfEE, Education and Training Statistics, (1997, 1998 and 1999 editions)

It should be noted that most recent developments in the provision of government assisted training have taken place in an uncertain job market for school leavers impacted upon by widespread technological change, the withdrawal of social security benefits for most youths aged under 18 years. Benefits were withdrawn for this age group following the implementation of the Social Security Act 1988, because they were guaranteed a place on the youth training programmes.[77] However, to facilitate movement between YT programmes, an eight-week bridging allowance was subsequently introduced (Bradley, 1995, p. 32).

Moreover, while the various programmes have been designed to subsidise youth employment and encourage training with minimal disruption to the youth labour market, their introduction has fundamentally altered this market. The plethora of training and education reforms has transformed the school-to-work transition by increasing the range of choices available to 16 year olds. For instance, Micklewright (1989) examines different determinants affecting the probability of leaving school at the minimum school leaving age using data from the National Child Development Study (NCDS). Family background, as measured by parental education, class, and number of siblings, was found to have a substantial impact on leaving school at 16.[78] Micklewright's analysis focuses on the early 1970s, a period he describes as "...more or less free from demand-side constraints and specific government intervention in the youth labour market" (p. 27). The NCDS respondents became eligible to leave school in 1974 and would have been too old to qualify for the YOP in 1978. In contrast, Andrews and Bradley (1997) draw on Lancashire Careers Service records to model the school-to-work transition (and demand for training) for a cohort of 16-year olds in 1991, a period in which they faced a wider set of choices. To reflect the institutional reform that had taken place in the l980s, Andrews and Bradley modelled six choices/outcomes: non-vocational continuing education, vocational continuing education, youth training, employment with on-the-job training, employment with general skills training, and unemployment. They model the first destination of the respondents, 6 months after becoming eligible to leave school, using a Multinomial Logit. The study found that, in addition to family background, local labour market conditions and schooling variables (including size and academic performance of the school) significantly impacted on the school leaving decision.

---

[77] Robinson (1996) discusses benefit and other recent passive labour market policy reforms in more detail, including changes in benefit allowances, eligibility and the legislative framework.

[78] Family background factors were also identified by Dearden and Heath (1996) as the key factors determining staying-on rates in Australia even after the introduction of AUSTUDY, an education allowance payable directly to students from relatively poor backgrounds, in 1987.

In summary, each of the youth training programmes introduced in the UK has had distinct objectives, reflecting the economic and political circumstances at the time of their inception. Yet, they are clearly linked and can be usefully thought of as different generations of an evolving training scheme, which began, with the introduction of the YOP in 1978 (Bradley, 1995, pp. 33-41). Adopting a similar mode, Deakin (1996, p. 80) views the different evolutions of intervention as a "...policy-learning process". Over the last three decades, the emphasis of training policy in the UK has clearly shifted from regulation of employers to subsidies for the training of young people and the unemployed (Stevens, 1999, p. 17). The bulk of training delivery has evolved from 'learning by doing' structure, implicit in the work experience element of YOP, to the formal classroom learning inherent in later programmes. Similarly, over time, nationally standardised qualifications have become an integral part of the programmes explicitly recognising the learning that takes place. Specific programme objectives have however, at times been ambiguous, but the primary aim, common to all the programmes from the outset, remains the successful transition from school-to-work (thereby combating youth unemployment).[79] The programmes have addressed many of the rigidities inherent in the traditional apprenticeship system but, as noted by Payne (1999), they have had difficulty shaking off the image that they are primarily designed for unemployed youth who have few other options. The increased emphasis on flexibility and progressive qualifications in the Modern Apprenticeships, Traineeships/Skillseekers system, combined with the launch of the employment (manpower) policy orientated New Deal to operate in parallel with the training programmes, should help address this image.

4.3: Evaluation in the United Kingdom

In an age of constrained budgets and increased accountability, evaluation of social programmes plays an integral part in programme monitoring and development.

[79] More recently the D*f*EE has attempted to clarify the aims of employment policy by adopting the term 'employability' as a framework for policy analysis. In a recent Department Research Report, Hillage and Pollard (1998) define employability as "...having the capability to gain initial employment, maintain employment and obtain new employment if required" (p. 1).

Moreover, the justification and success or otherwise of government intervention in the youth training market relative to other sorts of provision is subject to considerable debate. For instance, considerable levels of public funding have been injected into the various programmes, which over the period 1978 to 1993 benefited in excess of 5 million young people. Dolton (1993, pp. 1266-1270) estimated that within this time period funding (in 1990 prices) ranged from £1200 (1978/79) to as much as £3000 (1990/91) per trainee. Although it has since tended to fall in real terms the cost of training placements greatly exceeded the cost of having that person on the unemployment register.[80] In a recent *News Update* the Department for Education and Employment (D*f*EE) estimated the 'true' cost of the jobs secured in the first 6 months of the New Deal was around £1,000 per job. The New Deal was, however, only in its infancy at this time with most entrants still in the gateway stage of the programme, as such this is likely to be a conservative estimate of the true cost. Moreover, a number of employers have recently complained about the high bureaucratic cost in implementing the New Deal. The difficulty that arises in trying to apportion the state subsidy to individual participants and employers is discussed in Dolton (1993, pp. 1269-1270). For example, an investigation by the National Audit Office (NAO) into the performance of Skillseekers suggested that employers contributed £1.53 for every £1 spent by the Scottish Executive but it could not establish if the training would have taken place in the absence of the programme. Expenditure allocated directly to Skillseekers in 1998/99 amounted to approximately £69 million. While the overall UK figures for expenditure on work-based training for youth in 1998/99 amounted to £741 million.

Evaluation can provide useful feedback to policy makers in determining if this money is being effectively spent. However, as discussed in chapter 3, economic analysis is complicated by sample-selection problems, and problems associated with disentangling the effects of training from other factors. This difficulty is further

---

[80] Central government expenditure on youth training is mostly captured in the ONS classification *work-based training for young people*. In 1990/91 expenditure on this measure stood at £869 million compared to £723 million in 1998/99 (both measures expressed in 1997/98 prices). Over the same time period annual expenditure in Scotland on Skillseekers has been reduced from £90 million to £79.5 million (NAO, 2000, p. 14).

exacerbated by the practical difficulty of defining, and hence measuring, exactly what constitutes training. On-the-job training structures typically range from unstructured 'learning by doing' to formal apprenticeship arrangements, while off-the-job-training structures increasingly overlap with traditional education routes. Moreover, the diverse nature of youth training programmes further complicates the definition of training. For example, some companies use the programmes as a basis for their apprenticeship training while others use them to screen potential employees.[81] While additional programme placements cater for people with special needs, programme delivery is segmented (different modes) and in general, programme entrants are those young people with the poorest educational background (Robinson, 1996, p. 80). Equally, subsidies to private-sector training make it difficult to distinguish public sector training from private sector training and vice versa. For example, the Scottish Audit office estimates that on average employers spend £1.53 on Skillseekers for every £1 of government expenditure. More significantly, from an evaluation perspective, programme objectives are often ambiguous, multiplicative and outcomes are not always readily quantifiable.

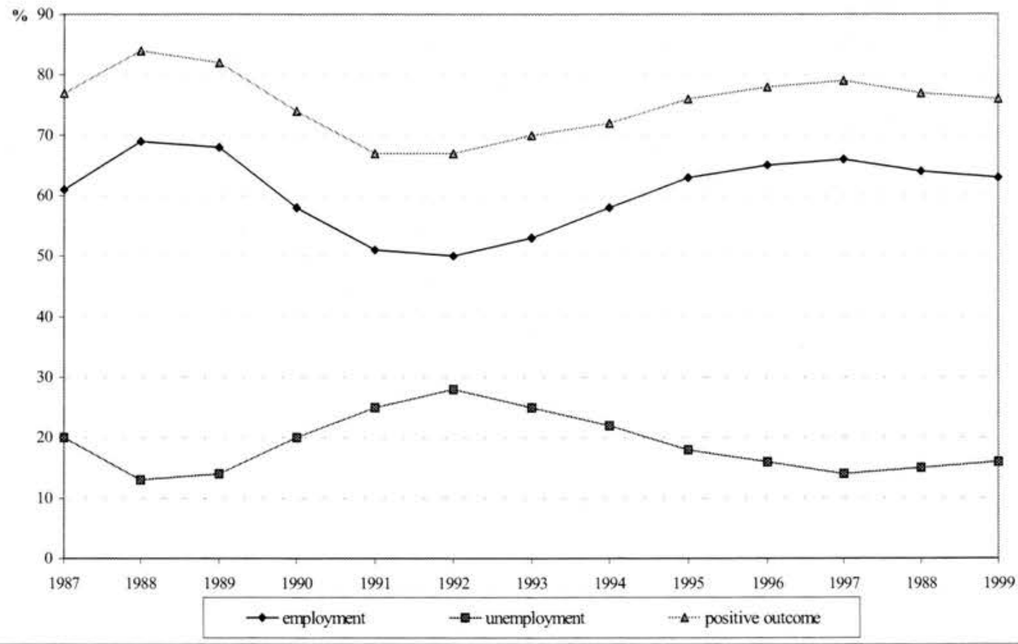
Early approaches to training programme evaluation involved cost-benefit analysis based on aggregate measures relating to the relative employment probabilities and mean post-programme earnings of participants and non-participants (Dolton, 1993, pp. 1270). This approach does not allow the adequate assessment of the effect of training on an individual's earnings or employment prospects while controlling for interpersonal differences. Nor does it adequately deal with problems arising from sample selection. More frequently, success or failure was judged simply by recording the outcome of the programme in terms of jobs obtained (placement rates). Deakin (1996, p. 79) notes that under this measure the YOP can be considered a partial success having helped approximately 60 percent of participants into

---

[81] Sako and Dore (1986) and Begg, Blake and Deakin (1991) provide empirical evidence, based on employer surveys, that employers place most emphasis on the value of YTS in selecting employees with the greatest potential. This implies that they undervalue government assisted training programmes in terms of skill enhancement (increasing human capital).

employment.[82] Post-programme destinations of young people, six months after completion, continue to be used by the D*f*EE for monitoring programme performance. Figure 4.2 illustrates reported post-programme destinations of participants between 1987 and 1999, in terms of being in employment, unemployment or a 'positive outcome'. The D*f*EE defines a positive outcome as being in work, education or entering a further training programme. More recently, this form of monitoring has been complemented with the setting of national targets for education and training, in terms of qualifications gained, and progress toward them is reported annually (Payne, 1999, pp. 484-485).

Figure 4.2: Destinations following youth training in the United Kingdom, 1987-1999



Source: ONS, Labour Market Trends (various years)

In Scotland monitoring of Skillseekers has indicated that outcomes, in terms of placement rates, have remained relatively stable over the period 1995/96 to 1998/99.

[82] Deakin (1996, chapter 5) reviews a number of assessments of early interventions in the UK youth labour market.

Three months after completing their training, approximately 56 percent of respondents to follow-up surveys claimed they were in employment or self-employed and a further 14 percent in further training. In other words, a positive outcome is observed in approximately 70 percent of cases (NAO, 2000, pp. 31-32).[83] Alternatively, using the same methodology, vocational qualifications successfully attained by trainees can be used as the outcome measure. For example, the National Audit Office report on Skillseekers found that over time the LECs network had achieved qualification targets in line with expectations. The report found that the number of trainees attaining level 3 qualifications fell short of targets but this was compensated by over achievement in level 2 targets (ibid. p. 28).

It is worth noting that a potentially negative feature of this type of programme monitoring is that the requirement to meet numerical performance targets for political as well as policy reasons can easily take precedence over the needs of the individual. Microeconomic evaluation, on the other hand, attempts to establish what type of individual derives benefit from the relevant training programme and by how much. Evaluation based on more sophisticated econometric techniques, sometimes called 'second generation studies', only became possible with the advent of large-scale individual cross-section and longitudinal data collection in the 1970s (Dolton, 1993, p. 1271). This allowed attention to shift to the net impact of training on the individual participant and hence allow the relative effectiveness of the programme to be assessed.

From a human capital perspective, the success of labour market training is ultimately determined by the magnitude of enhancement of an individual's lifetime earnings, however at the time we are required (or choose) to evaluate the programme, we have, at best, an incomplete picture of lifetime earnings. Moreover, as established in earlier chapters, we cannot observe the counterfactual. That is what actual earnings would have been in the absence of training. The evaluator thus needs to choose an appropriate outcome of interest as a proxy for lifetime earnings and model the

---

[83] Since September 1999 the surveys are being conducted 6 months after training making future comparison with England more feasible.

counterfactual to approximate the true impact of training in terms of the chosen outcome. Chapter 3 provided a range of methodologies employed by economists to model training decisions and programme effect and most of these are illustrated in this section in which the UK evaluation literature is reviewed including polychotomous choice and duration models. Using these techniques it is possible to derive a range of different outcomes of interest to analysts and policymakers, however, the bulk of empirical research has focused on just two key indicators. These are earnings and employment probability subject to participation in the relevant programme.[84] Clearly there are many other criteria on which to judge the success or failure of a programme, such as the quality of any jobs obtained, qualifications gained and social welfare considerations, all of which are required before a comprehensive cost-benefit analysis is possible. Attempts to assess programme impact using such outcome measures are typically limited by a dearth of relevant data. Assessing the impact of the programme on the individual, while controlling for interpersonal differences, is however a significant step toward such comprehensive evaluations.

Given the emphasis of second-generation studies on earnings and employment, the findings reported in the relevant literature are appraised separately for each outcome below. This is preceded by a brief examination of the datasets that the majority of these estimates rest upon and followed by a discussion of qualifications attained as an alternative outcome measure. Section 4.4 completes the analysis of UK evaluation activity by summarising the employment studies on the basis of model choice and specification to highlight the possible impact of decisions made by the researcher on resulting impact estimates. Institutional developments and evaluation activity relating to the provision of youth training in Australia are then examined, commencing in section 4.5.

4.3.1: Data Sources in the United Kingdom

All of the estimates from the UK evaluations reported in this chapter have been generated using non-experimental data. Moreover, most of the studies have utilised

---

[84] Also see the discussion in section 2.4 (chapter 2 of this thesis).

the two complementary government-sponsored surveys: the Youth Cohort Study (YCS), covering England and Wales; or the Scottish Young Person's Survey (SYPS) covering Scotland.[85] These surveys gather comprehensive data on different cohorts of the 16 to 19-year age group, collating data from the respondent's schools and other educational authorities in addition to the respondents themselves. The central feature of these surveys is a diary, (administered by postal questionnaire) which cohort members complete, indicating their labour market or educational status for regular periods of the survey. This allows the construction of a longitudinal record of their labour market status. The aim is to collect information on all education, training, jobs and periods of job search the individual had in the relevant period. In addition, the surveys also gather a wide range of socio-demographic detail allowing characteristics, such as family background, to be included in any evaluation of the data. Some differences have developed between the two surveys over later sweeps but the vast majority of characteristics are shared.

A variety of different factors impact on the validity of the survey datasets for evaluation purposes with respect to training. The use of the retrospective diary, as opposed to personal interviewing, means that the data are largely based on the young person's own reports of their full-time status and as such may generate recall problems and inaccurate recording (unintentional or otherwise). Recall problems may be particularly acute with respect to events towards the beginning of the recall period. It is also possible that observations only capture the main activity as reported by the young person and may not accurately reflect the dynamic nature of the youth labour market. For example, the respondent may have changed jobs, moved in and out of training programmes, left school or made some other transition within a given period between reporting times. The surveys are also subject to attrition, which gives rise to a potential source of bias if the attritors differ substantively from those who remain in the sample. Further, while the apparent richness of the datasets make them attractive for the analysis of training, they are not solely gathered for this purpose

[85] The 4th cohort of the SYPS is used below, in chapter 6, to evaluate recent returns to youth training in Scotland. A more detailed description of the dataset is hence provided in section 6.4.

and have to serve a wider research community. In particular, it is difficult to determine differences in training quality or even participation in different categories (modes) of the programmes. Moreover, a major limitation of the data arises from the survey's concentration on the 16 to 19-year age group, when it has become increasingly common for training and education to extend beyond this period.[86] As such, the range of possible specifications is limited and the analyst needs to allow for the potential bias inherent in the survey data. The constraint imposed on evaluation by data availability in the UK has led Dolton, Makepeace and Treble (1994c) to observe that:

> "To a great extent, our problems are dependent on the nature of the data which we are working, and it is sometimes difficult to distinguish a truly conceptual issue from one that arises simply from the structure [of the dataset]" (p. 195).

A number of studies have used an alternative dataset, the National Child Development Study (NCDS). It has predominately been used to estimate the impact of job related training in the private sector, for example see Baker (1990) or Blundell, Dearden and Meghir (1996). It has also been used to analyse the school to-work transition in terms of the school leaving decision (see Micklewright, 1989). The NCDS is a nationally representative longitudinal study of 18,000 individuals born in 1958. In the fourth sweep (NCDS4), conducted in 1981, just over 12,000 individuals remained in the sample. The study contains a rich source of information on the respondents work history since leaving school allowing longer term consequences of early training upon earnings, earnings growth and employment to be undertaken. The NCDS has provided useful feedback on the performance of employer-led training such as apprenticeships but the cohort respondents were already 20 years of age in 1978 and hence ineligible for any of the large-scale programmes that began with the introduction of the YOP in that year. Indeed only 4 percent of the NCDS cohort who left school in 1974 was unemployed in January 1975 compared with 22 percent of 16-year old school leavers a decade later

---

[86] An exception is the third cohort of the England and Wales Youth Cohort Survey, which re-surveyed its cohort when they reached approximately 23 years of age. Payne (1995) has used this cohort for training programme evaluation; see sections 4.3.2 and 4.3.3 below.

(Micklewright, 1989, p. 27). The British Cohort Study is a similar study of individuals born in 1970 and thus eligible for YTS. Green, Hoskins and Montgomery (1996) use a sweep of interviews conducted between January and April 1992, when the respondents were aged 21, to estimate the earnings impact of taking different routes in the school-to-work transition (see discussion below in section 4.3.2).

Various government departments and agencies also compile data that can be used directly or in conjunction with other databases for training evaluation. For example, the Careers Service collects data on a regional basis generating a wide range of information on the personal, educational and, socio-economic characteristics of young people in delivering its statutory duties of vocational guidance and a placement service (Andrews and Bradley, 1997, p. 394). The database constructed by the Lancashire Careers Service, which has again been used for the evaluation of training in the private sector and analysis of the school leaving choice, tracks individuals for a number of years enabling different choices in the school-to-work transition and in occupation. In December 1991 the Lancashire Careers Service, in conjunction with the local TECs, also surveyed employed school leavers in 1991 to assess the type and level of training young people were receiving in the region thereby allowing analysis to be extended to training. Other examples include the Scottish School Leaver's Survey and the National Labour Force Survey. Moreover, in addition to regulation stipulating that each member country must include questions within its Labour Force Survey that elicit information on a range of employment and unemployment issues, the European Union requires monitoring and evaluation as a condition of funding for employment related projects.

4.3.2: The Impact of Training on Earnings

The earliest second-generation studies rested heavily on the human capital notion that training is an investment leading to what Dolton (1993) describes as "...a preoccupation with wages" (p. 1271). For example, when evaluating early American interventions Ashenfelter (1978) concentrated on calculating the rate of return on training taking into account any forgone earnings. Subsequent studies (Bassi, 1984; Ashenfelter and Card, 1985) tended to concentrate on establishing if participants gained a wage premium relative to non-participants. Moreover, enhancing the earnings of trainees was the principal objective (and hence indicator) of government intervention in the USA where it was also sometimes perceived as a measure of the quality of the training (Bassi, 1984). Consequently, the evaluation of training programme effectiveness in terms of estimated earnings impact has, to a great extent, become the industry standard.

Reported earnings impact estimates of government assisted youth training programmes in the UK are summarised in table 4.1 below. Authorship, listed in column one, and the relevant database, in column three, identifies each study. The second and fourth columns of the table illustrate the variety of possible specifications (the key to the variables common to different studies is given at the base of the table). It can clearly be seen that, over time, analysts have utilised increasingly sophisticated models and employed a greater variety of descriptive variables. In general, however the analyses are all based on the earnings equation introduced by Mincer (1974) which proposed that, for each individual, the natural logarithm of earnings systematically depend on the type and amount of training, individual characteristics and various background characteristics. Consistent with the human capital theory framework, the emphasis is mainly on individual characteristics that might dictate investment in training, such as educational attainment and family background. The fifth column shows the wide range of associated impact estimates. It is readily apparent that most studies show a negative impact on the earnings of programme participants relative to non-participants. Given the duration of most of the surveys, this is consistent with human capital theory, which suggests that, during investments in training, wages will be lower than otherwise. This implies that participation in training will be associated with lower initial earnings compared to those of non-participants with the same ability. It is expected, however, that the investment in training will be followed by higher wages.

Hutchinson and Church (1989) used a 1984/85 survey of firms in five local labour markets (Reading, Preston, Torquay, Motherwell and the London Borough of Newham), administered by the MSC to monitor wider labour market effect of interventions, to evaluate the early impact of the YTS and YWS. They found that the wages of 16 year-olds and adults (over 18 years) that participated in the YTS were respectively 8.2 percent and 7 percent lower than for non-participants in the same age groups but evidence of a wage effect for 17 and 18 year olds was absent (statistically insignificant). On examining the effectiveness of the YWS, Hutchinson and Church found that, to a large extent, it was successful in lowering the wages of 17 and 18 year olds. This may have impacted on some of the ex-YTS participants, however relatively few firms took part in both schemes. They also estimated the impact of unionism on youth earnings estimating a small positive gain from membership (6.5 percent for 18 year olds) that increased with age (8.5 percent for adults).[87] Hutchinson and Church employed OLS to estimate the relevant wage equations in their analysis but they did not allow for possible sample selection bias.

Although not estimating the earnings impact of YTS directly, a study by Main (1987) provides an interesting contrast to the Hutchinson and Church study. Main adopted a different approach to assess the merit of the YWS and YTS by comparing the expected earnings of a group of unemployed school-leavers with the actual earnings of a group of employed school-leavers. This framework allowed him to determine if the underlying assumption of the YWS (and its successor the NWS) was that wage expectations of young job seekers were unrealistic relative to market rates. His analysis showed that, after allowing for possible sample selection, it was possible to reject the hypothesis that expected earnings of the unemployed were consistent with the observed market earnings of employed school-leavers. For most classifications tested, expected earnings of the unemployed were found to be below earnings of the employed. With respect to YTS, the expected earnings of participants exceeded those of the unemployed but still remained below observed market earnings of the employed school-leavers.

Main and Shelly (1990) used SYPS data on 1983/84 school leavers to analyse the affect of the one-year YTS (YTS-1) on participant wages in 1986 (approximately 18 months after leaving school). Their analysis specifically allowed for the possibilities

---

[87] Comparative research by Tan, Chapman, Peterson and Booth (1991) found that among non-apprentices, union members were more likely to receive training in both Australia and the UK. In the UK, unionised apprentices were also more likely to receive off the job training but this relationship did not hold for Australia. For an analysis of union attitudes and policies toward the YTS see Ryan (1995).

of sample selection by applying Heckman's two-step correction model, using a Probit equation to determine YTS participation. A positive but statistically insignificant impact on subsequent earnings is identified for 'disadvantaged' school-leavers.[88] The estimated impact for an 'advantaged' school-leaver is negative but again statistically insignificant and, as such, Main and Shelly stress "...caution against drawing any definite conclusions from them" (p. 509). Nevertheless, the study does illustrate two distinct questions that can be asked when comparing counterfactual outcomes. The first asks, "What would be the impact of training on earnings if people were randomly assigned to training?" And the second question asks, "How do the post-programme earnings of the trained compare to what they would have been in the absence of training?"[89] Main and Shelly show that YTS has a negative wage impact estimate for advantaged school-leavers when the subject is randomly assigned to training. In contrast, when examining the wage impact on a YTS participant (the impact of 'training on the trained') the estimates are positive. The difference is largely due to different estimates of market worth without YTS, indicating that advantaged YTS participants are different from the average advantaged school-leaver in ways that are not explicitly measured in the data. This is consistent with the observation made by Heckman and Robb (1985, p. 161) that the two questions will only have the same answer when training has the same impact on everyone or there is random assignment to training.

[88] Main and Shelly (1990) define a disadvantaged school leaver as: male, aged less than 18 years, unqualified, with an unemployed father and in an area with an unemployment rate exceeding 21.1 per cent (p. 501).

[89] This approach was first advocated by Heckman and Robb (1985, p. 161); also see discussion in section 3.2 (chapter 3 of this thesis).

Table 4.1: Summary of studies on Youth Training in the UK by estimated impact on earnings

Author	*Method*	*Database*	*Explanatory variables*[1]	*Earnings effect*
Main (1987)	Analyses expected and actual earnings of school leavers, including those opting for YTS.	1985 Scottish School Leaver's Survey eligible for YOP	A, B, H plus expected earnings dummy	Expected earnings of YTS exceed those of the unemployed
Hutchinson & Church (1989)	Attempts to estimate the earnings impact of YTS, YWS and unions using OLS.	1985 YOPELM YOP/YTS-1	A, B, H plus industry, occupation and union status	16-year olds: -8.2%* 17-year olds: -3.6% 18-year olds: +0.04% Adults: -7%*
Main & Shelly (1990)	Attempts to estimate impact of YTS on hourly earnings allowing for sample selection using a participation equation then applying OLS	SYPS, 1983/84 school leavers YTS-1	A, B, H, J plus completion of YTS and dummy for remaining with YTS employer (YTS job)	*Disadvantaged groups* YTS Employer: +22% Non-YTS : +28% *Advantaged groups* YTS Employer: -8.3% Non-YTS : -4.4%
Whitfield & Bourlakis (1991)	Similar to Main & Shelley but applied to YCS. OLS reported but also estimated Probit/OLS	YCS, 1983/84 school leavers YTS-1	A, B, H, J plus ethnicity and YTS job	*Average* -3% YTS Employer: -8%
Dolton, Makepeace & Treble (1994a)	Aimed to replicate specifications of earlier studies but with later YCS cohort	YCS 1985/86 school leavers YTS-2	A, B, HV, J plus, transport, housing tenure, & duration of unemployment.	*Average* Female: -4.5%* Male: +4.7% *Non-YTS > YTS* Females: 9 to 10% Males: 5 to 6%
Dolton, Makepeace & Treble (1994b)	Multinomial Logit: 0=do nothing 1=some FE, not YTS 2=no training but YTS 3=off-the-job, not YTS 4=off-the-job and YTS 5=non-YTS apprentice 6=YTS apprentice.	YCS 1985/86 school leavers YTS-2	A, B, HV, J plus transport, literacy, numeracy, housing tenure & duration of unemployment focus on private and public sector provision	*Average* Female: -16% Male: -15% (c.f. do nothing) *Apprentice* Non-YTS > YTS (-11% cf. -27%)
Payne (1995)	Multinomial Logit: 0=full-time higher Ed. 1=full-time vocational Ed. 2=full-time mixed Ed. 3=apprenticeship. 4=non-apprentice YTS	YCS 1985/86 school leavers YTS-2 At 23 years	A, B, HV, J plus any disability, housing tenure Also allow for time spent on education-training route	Female: -15% Male: -8% Male apprentice: +4%
Green, Hoskins & Montgomery (1996).	Multinomial Logit: 0=do nothing 1=some FE, no training 2=FE and training 3=training, work at 16 4= YTS & qualification 5= YTS & no qual. 6=any YTS.	1970 British Cohort Study YTS-2 At 21 years	A, B, HV, J plus any disability, literacy, numeracy, specific skills, responsibility & family firm.	-15% (3 years after YTS)

[1.] *Key: A = Personal characteristics: age, gender, marital status*
B = Background characteristics: parental occupation, father's social class, family size, region
H = Human capital measures: educational attainment
HV = as for H plus measures of vocational educational attainment
J = Job characteristics: regional unemployment and wage, firm size, occupation and/or industry
* *Reported as statistically significant at the 5% level*

Whitfield and Bourlakis (1991), adopted similar specifications and method to Main and Shelly in order to estimate the earnings impact of the one-year YTS on 1983/84 school leavers in England and Wales using YCS data. They estimated an average reduction of approximately 3 percent in the earnings of participants in the one-year YTS two years after completion. The estimated impact of a participant remaining in employment with their YTS sponsor was an 8 percent reduction in earnings. Whitfield and Bourlakis derived their reported estimates using orthodox regression methods but they also estimated the equations using Heckman's two-step method reporting that it made no material difference to the estimates. Dolton, Makepeace and Treble (1994a) found the negative impact persisted with the introduction of the two year YTS programme (YTS-2), at least in the case of women, estimating that YTS lowered female earnings by 4.5 percent. On the other hand, male earnings showed an improvement of 4.7 percent, although the estimate was not statistically significant. The authors also estimated the earnings by gender with non-YTS off-the-job training experience and found these to be significantly higher than their peers who had participated in YTS - a difference of 9-10 percent for women and 5-6 percent for men. The negative results from these earlier studies are not altogether surprising when consideration is given to the initial objectives of intervention in the UK. As observed in section 4.2, one of the aims of the MSC was to suppress youth wages in order to make young people more attractive to employers. This aim became more explicit with the introduction of the YWS and NWS.

The general approach in these earlier studies was to model the training option as a dichotomous choice comparing the set of all training programme participants against the set of all non-participants in that programme. Partially motivated by the rapid transformations of the youth labour market and aided by access to appropriate databases and modelling developments, more recent studies have tended to model the training option as a polychotomous choice. This approach better reflects the different possible school-to-work transition routes school-leavers face in the attempt to satisfy their occupational ambitions. Dolton, et al., (1994b) distinguish several alternative routes in the school-to-work transition, and estimate earnings for each route after allowing for selection into each. The negative impacts persist, however, with a male apprentice not only earning less (27 percent) compared to a typical individual with no training but also less (11 percent) than a fellow apprentice who opts without the YTS. The authors also stress significant differences between those that complete their training courses and those that abandon them part way through.

Green, Hoskins and Montgomery (1996) adopt similar methodology and specifications to Dolton, et al., (1994b) making use of a dataset compiled from a 10 percent sweep of the 1970 British Cohort Study when cohort members reached 21 years of age in 1991. In spite of the extended period of observation their results tend to concur with the earlier studies that YTS participants were still likely to be earning less (15 percent) than those with no training three years after completion. Payne (1995) makes use of an additional follow-up sweep in the third cohort (those legally entitled to leave school in the summer of 1986) of the YCS when the respondents reached 23 years of age. This introduced the opportunity to assess the performance of YTS relative to higher education and apprenticeships, which would not normally be completed in the period of observation covered by the other surveys. Compared to the higher education route, male YTS participants earned 8 percent less and females 15 percent less. For men, the apprenticeship route appeared to be quite favourable, with Payne estimating a 4 percent gain in earnings at 23 years. On the other hand, women opting for the apprenticeship route incurred an estimated 17 percent reduction, although not statistically significant.

A number of factors contribute to the predominately negative estimates of earnings impact observed in the UK evaluation literature. A key contributory factor is that most studies are constrained to observing only short-term post-programme earnings. Combined with the point-in-time analysis, inherent in the standard model, this generates post-training snapshots early in the individual's life-cycle earnings. As such, Bradley (1995, p. 47) suggests that it may be better to focus on the rate of growth in earnings although this would place greatly increased demands on data collection. Another notable factor is that many participants enter further training programmes (or higher education) on leaving YTS. While the Department for Education and Employment may consider this a 'positive result', in the short-term it will result in relatively lower evaluation payoffs, in either terms of either earnings or employment impact. Moreover, the need to consider the policy context, in terms of programme objectives and availability of alternatives, is highlighted when interpreting estimated results of the impact of youth training on earnings in the UK, illustrated by the co-existence of the YTS and YWS. The estimates are also likely to be biased downwards because training impacts on both employment and wage levels and if the programme is successful in the former it will increase the supply of relatively skilled youth labour thus reducing pressure on wages. To illustrate that the variation in reported non-experimental estimates of programme impact are not confined to the UK and to allow comparison with experimental estimates, we now look a selection of evaluation results from the USA.[90]

Comparison with Evaluations in the United States of America

In 1985 the United States Department of Labor funded a methodological conference on the reliability of non-experimental estimation for the evaluation of manpower training programmes. This included the funding of the National Supported Work Demonstration experiment (NSW), which included a randomly assigned control group. The immediate impetus for the conference was a series of studies on the incumbent training programme that presented startlingly different empirical estimates for the impact of training (see Stromsdorfer, 1987). The Comprehensive Employment and Training Act (CETA) governed programmes in place at the time. Evaluation estimates of the impact of CETA programmes on earnings were widely divergent. In fact, depending on the particular study chosen, one could conclude that CETA programmes were either quite effective in improving post-programme earnings or, alternatively, that the same programmes reduced the post-programme earnings of participants relative to non-participants.[91] The reported estimates from a

---

[90] Tan et al., (1991) observe that the earnings impact of training is much higher in the USA than in either the UK or Australia (roughly double) which is likely to impact on the incentive for workers to get and employers to provide training.

[91] The outcome of interest in the CETA programmes, introduced in 1973 was the wage, given the explicit purpose of the Act, as reported by Bassi (1984, p. 36), was "...to provide training and employment opportunities for economically disadvantaged, unemployed persons which will result in an increase in their earned income".

selection of non-experimental studies that specifically focused on youth participants in CETA and similar programmes are summarised in the upper portion of table 4.2. For a comprehensive review of the CETA evaluations see Barnow (1987).

Table 4.2: Summary of studies on Youth Training in the US by estimated impact

Earnings are reported as increase in post-programme earnings

Non-experimental Estimates: [1]

Author	*Training Cohort*		*Earnings: Males ($)* [a]	*Earnings: Females ($)* [a]
Cooley et al. (1979)	1969-71 MDTA		1, 230	600
Gay & Borus (1980)	1969-72 Job Corps	*whites* *minorities*	-215 148	-1, 282 -325
Maller et al. (1982)	1977 Job Corps		1, 600	800
Dickinson et al. (1986)	1976 CETA		-1, 100	370
Bryant & Rupp (1987)	1976 CETA		60 [b]	
Bryant & Rupp (1987)	1977 CETA		1,050 [b]	
Bassi et al. (1984)	1977 CETA	*whites* *minorities*	-1, 010 -1, 330	80 260

Comparative Experimental Estimates: [2]

		Control Group (N = 566) [c]	*Comparison Group (N = 566)* [c]
Fraker & Maynard (1987)	1977 1978 1979	313 -28 -18	-668 -1, 191 -1, 179

Sources: [1] *LaLonde (1995), adapted from Table 1, p.157.*
[2] *Fraker & Maynard (1987), adapted from Table 1, p. 98*
[a] *All earnings are expressed in 1990 dollars*
[b] *Separate estimates for males and females were unavailable*
[c] *Data are based on the National Supported Work Demonstration and the Current Population Survey*
Note: The Manpower Development and Training Act (MDTA) was the immediate forerunner to the CETA programmes. Job Corps came under the auspices of the MDTA

Influential studies by LaLonde (1986), LaLonde and Maynard (1987) and Fraker and Maynard (1987) set out to assess the validity of the non-experimental estimation techniques by comparing the corrected estimates of regression parameters, obtained from a non-experimental design with those derived from a randomised field experiment constructed using NSW data. Their findings showed that non-experimental techniques, on the whole, failed to replicate the experiment's results and were vulnerable to specification errors. As discussed in chapter 3, this means of evaluating non-experimental estimates is not problem (or assumption) free and subsequent work using the same database by Heckman, Hotz and Dabos (1987) came to the opposite conclusion. Moreover, experimental estimates do not entirely mitigate the divergence in estimated impacts. The randomly selected control group estimates, together with those generated by econometric methods (comparison group) reported by Fraker and Maynard, are summarised in the lower section of table 4.2.

4.3.3: The Impact of Training on Employment

The estimates of the impact of youth training programmes in terms of enhancing employment prospects, subject to participation, are summarised in table 4.3, which utilises the same format as used in the earnings impact summary reported in table 4.1. Column 2 clearly illustrates the extensive range of different methodologies employed to evaluate the programmes. For example, Main and Raffe (1983) use the standard OLS; Main (1985) employs a Probit equation, Payne (1995) a Multinomial Logit and Mealli, Pudney and Thomas (1996) duration analysis. The fifth column highlights the considerable variance, albeit generally positive, which is evident in reported impact estimates of training in terms of the conditional probability of gaining employment. For instance, in their analysis of the one-year YTS using YCS data on 1983/84 school-leavers, Whitfield and Bourlakis (1990) estimated a 4 percent increase in the probability of YTS participants being in employment at 19 years of age. However, this was insufficient to offset the negative effects associated with the local unemployment rate, disability or ethnic origin (pp. 48-50). In contrast, in their analysis of SYPS data on 1983/84 school-leavers, Main and Shelly (1990) estimated a positive impact of between 11 and 17 percent, dependent on the type of school-leaver. The latter figure refers to the group defined by the authors as advantaged school leavers. Main (1991) reported an even larger increase for the same group (14-19 percent) using the same data set but utilising observations on the respondent's labour market status six months further on.

Employing a similar model to these earlier studies and drawing on the YCS (also 1983/84 school leavers), O'Higgins (1994) produced similar (averaged) estimates to Main and Shelly. However, after correcting for heteroskedasticity and employing a bivariate switching regression model (as an alternative to the selectivity model), O'Higgins estimated the impact of the YTS on participants one year after completion as falling in the range between 1 and 21 percent, dependent on the type of individual. Dolton et al. (1994a) also attempted to replicate the results of the two earlier studies by using similar model specifications but this time used later YCS data (1985/86 school leavers) allowing the two-year YTS to be assessed. They identified a slight negative effect in general however they do report a positive effect for males in certain specifications, particularly in areas of high regional unemployment.

Using the additional sweep of the third cohort of the YCS, Payne (1995) modelled the risk of unemployment at 23 years of age over a selection of paths in the school-to-work transition, including those who chose to stay on in full-time education. She found that participants in the non-apprenticeship YTS and apprenticeship routes were more likely to be in employment at 23 years than those who pursued either the vocational or academic education routes when they were 16 years of age.

Dolton, et al. (1994a) applied duration analysis to assess the effect of the YTS on the probability of gaining employment over time, rather than at a specific point in time. Focusing on the length of time it took to find 'good' or 'bad' jobs. To do this they constructed a work history between September 1986 and February 1989. Their analysis found that YTS reduced the amount of time it took for women to find a good job. The YTS did not, however, reduce the amount of time it took men to find a job. An interesting observation made by the authors was that, when the time on YTS was netted out, even male YTS graduates found good jobs as quickly as non-graduates. They suggested that employers do not always associate the YTS with low levels of quality or motivation (p. 647). Mealli, Pudney and Thomas (1996) also applied duration analysis and found that completing the two-year YTS significantly enhanced the participant's employment prospects.

Table 4.3: Summary of studies on Youth Training in the UK by Estimated Impact on Employment Probability

Author	*Method*	*Database*	*Explanatory variables*[1]	*Employment effect**
Main & Raffe (1983)	Used OLS to estimate the effects of YOP on the probability of finding employment	1979 Scottish School Leaver's Survey eligible for YOP	A, B, H plus local unemployment rate & truancy	Females: +14%* Males: +6%
Main (1985)	Used Probit maximum likelihood to estimate the effects of YOP	1981 Scottish School Leaver's Survey YOP	A, B, H plus local unemployment rate	Females: +8.1%* Males: +4.4%* *Disadvantaged youth* Females: +6.5% Males: +2.6% *Advantaged youth:* Females: +8.4% Males: +5.1%
Main & Shelly (1990)	Binominal Probit to estimate employment impact of YTS allowing for sample selection.	SYPS, 1983/84 school leavers YTS-1	A, B, H, J plus completion of YTS and YTS job.	*Disadvantaged male:* +11% *Advantaged male:* +17%
Whitfield & Bourlakis (1991)	Binominal Probit: similar to Main & Shelley but applied to YCS.	YCS, 1983/84 school leavers YTS-1	A, B, H, J plus ethnicity and YTS job	avg +3%
Main (1991)	Binominal Probit: similar to Main and Shelley but concentrates on early school leavers	SYPS 1983/84 school leavers (+ 6 months) YTS-1	A, B, HV; J plus YTS completion, YTS job, Christmas & unemployed leaver	*Disadvantaged male:* +14 *Advantaged male:* +19%
O'Higgins (1994)	Bivariate switching Probit, also replicated specifications of earlier studies	YCS 1983/84 school leavers YTS-1	A, B, HV, J plus ethnicity, transport, housing tenure & truancy	+1 to 21%* *Average:* +9% *Disadvantaged* +11%
Dolton, Makepeace & Treble (1992a)	Binominal Probit: replicated specifications of earlier studies with later YCS cohort	YCS 1985/86 school leavers YTS-2	A, B, HV, J plus disability, literacy, numeracy, specific skills, responsibility & family firm.	*Average:* < 0 Male: > 0
Dolton, Makepeace & Treble (1994c)	Duration analysis: estimated time to a 'good' job	YCS 1985/86 school leavers YTS-2	A, B, HV, J plus transport, literacy, numeracy, housing tenure & duration of any unemployment	Female: > 0 i.e. Found a 'good' job sooner
Payne (1995)	Multinomial Logit: 0=full-time higher Ed. 1=full-time vocational Ed. 2=full-time mixed Ed. 3=apprenticeship 4=non-apprentice YTS	YCS 1985/86 school leavers YTS-2 At 23 years	A, B, HV, J plus any disability, housing tenure Also allow for time spent on education-training route	Apprentice: > 0 (YTS & NON-YTS)
Mealli, Pudney & Thomas (1996)	Duration analysis	Lancashire Careers Service database YTS-2	A, HV plus any disability, unemployment spells & occupational variables	91.4% if YTS-2 was completed cf. 57.6% for incomplete spells

[1.] *Key: A = Personal characteristics: age, gender, marital status*
B = Background characteristics: parental occupation, father's social class, family size, region
H = Human capital measures: educational attainment
HV = as for H plus measures of vocational educational attainment
J = Job characteristics: regional unemployment and wage, firm size, occupation and/or industry
* *Reported as statistically significant at the 5% level*

In general, the evaluation literature demonstrated that participation in YT was associated with an increased probability of subsequent employment. The impact is however, relatively modest and not as significant as other factors, such as better school leaving qualifications, early labour market experience or other measures of labour market advantage.

4.3.4: The Impact of Training on Other Outcomes of Interest

Few studies have examined outcomes of interest other than earnings impact or employment prospects. An exception is Payne (1995), who assessed the performance of the two-year YTS in terms of NVQs gained. Using a binary (yes/no) response measure from the third YCS cohort, indicating if the respondent has obtained NVQ level 3 or not, she was able to apply a logistic regression model in order to estimate how the different options taken at 16 years affected qualifications at 23/24 years.[92] Given the relative complexity of interpreting the estimated odds inherent in the logistic model, Payne provides fitted probabilities of the estimates for different sets of characteristics. Generally (non-apprentice) YTS is shown to be relatively poor at increasing the probability of the participant gaining NVQ level 3 qualifications but this is not altogether surprising as level 2 was the objective set by the programme.[93] In contrast, those entering the full-time academic education route were almost certainly setting out to gain higher qualifications, that is outcomes in terms of qualifications are constrained by the aims of young people before embarking on the school-to-work transition (Payne, 1995, p. 6). Moreover, the different training routes in this study are compared relative to the full-time education

---

[92] The options are the same as for Payne's assessment of earnings impact and employment prospects, reported in tables 3.1 and 3.3 above. NVQ level 3 was partially chosen because the National Target's for learning state that by 2000, 60 percent of 21 year olds should have reached the equivalent of NVQ level 3 (Payne, 1995, p. 6).

[93] Jones (1988, pp. 63-67) reported that less than a quarter of trainees on the one-year YTS gained a qualification of any sort and that the vast majority of these were fairly basic in character, roughly corresponding to NVQ level 1. Plotting trends in numbers gaining qualifications he concluded that participation in YTS did lead to an increase in total numbers gaining a NVQ level 1 qualification, 'perhaps in the order of 50 percent'. However, no discernible trend was identified in terms of higher-level qualifications.

option and it does not include an option of not undertaking post compulsory education or training at all. Significantly, the full-time vocational education and apprenticeship options (within and without YTS) perform comparatively well, particularly for males with a high exam score at 16 years of age.

Steedman and Hawkins (1994) also investigate the relationship between NVQs and youth training but assess it from an institutional perspective. They investigate how the formal requirement of NVQs under YT impacted on the scope and standards of initial training in the building trades. The focus of the study is on the mathematical requirements in the industry, which the authors identified as receiving less emphasis under NVQ level 2 than under the requirements of the former City and Guilds certification. Drawing on national statistics of numbers gaining qualifications and the mathematics component in recognised training syllabuses and examinations, they conclude that NVQ level 2 is unsuitable for trainees in the building trades. Marsden and Ryan (1990) review the institutional context of YTS and its impact on youth labour market outcomes. They note that occupations that require costly skill development have been adversely affected by the decline in traditional apprenticeships and the failure to build the YTS into a widely accepted substitute. The introduction of Modern Apprenticeships may go some way to addressing their concerns.

4.4: Identifying the Source of Variation in Impact Estimates

Formal estimates of the earnings impact and employment prospects of youth training participants have been mainly concentrated on the predecessors to the current form of the programme, the YOP and YTS. A number of elements are common to both programmes and similar datasets have been used in the evaluations. Nevertheless, sections 4.3.2 and 4.3.3 illustrated the variability of estimated programme effect that can arise from the application of non-experimental methods even when applied to broadly similar programmes using seemingly similar datasets. The use of different regions, time periods, training duration, numbers of observations and cohort sweeps may account for some of the observed variance in the estimates. The remaining variation between studies will be due to the models themselves or specification of these models, for example if trainees have been differentiated on the basis of gender and/or advantage. Furthermore, column three in both tables 4.1 and 4.3 illustrates that a wide variety of different background variables have been employed in different studies and these differences in specification may result in different outcomes even where the same model is applied.

To highlight the impact of possible choices made by (or imposed upon)[94] the researcher on discrepancies in outcomes, table 4.4 presents a summary of the employment studies reviewed in section 4.3.3, excluding the duration analyses, by a selection of different methodological classifications. Further classifications could easily be incorporated but this selection concentrates on model structure, which varies considerably between studies. Work experience is included because a number of studies have emphasised the importance of early labour market contact on future employment prospects (Main and Shelly, 1990; O'Higgins, 1994). Firm size, on the other hand, is considered to be a major determinate in the employer's decision on whether or not to provide training (Whitfield and Bourlakis, 1991).

A careful examination of this table together with table 4.3 may suggest a number of possible explanations for the variability in outcome estimates including the statistical control for sample selection into training. However, it is difficult to discern any clear pattern between the methodological choices and outcomes. In order to gain a clearer picture of the possible reasons for between study variation a detailed statistical analysis is required to investigate the effect of these specification choices. A method that can be used for this purpose is meta-analysis. This is a technique that provides a formal procedure for detecting any statistically significant differences across studies, using the summary statistics of the original studies as its dataset, and highlights the impact of specification choices on the relevant outcome. Chapter 5 presents a meta-analysis of the evaluation studies reviewed above (sections 4.1 and 4.3). The meta-analytical method is however, limited by the quality of the original studies and does not examine the workings of the different models. Thus, in order to make any definitive statements with respect to the impact of employing different evaluation

[94] For example, choices open to the analyst may be constrained by the unavailability of reliable data, timing of the study, etc.

models, a detailed empirical analysis that varies a number of economic and statistical assumptions one at a time on a common database is required. This form of analysis is known as sensitivity analysis and is applied to the earnings impact of YT in Scotland in chapter 5.

Table 4.4: Methodological classification of Youth Training studies in the UK[1]

Author	*Model*	*No. of options for dependent variable*	*No. of explanatory variables[1]*	*Exogeneity of Experience*	*Exogeneity of firm size*	*Sample selection controls*
Main & Raffe (1983)	OLS	2	9	No	No	No
Main (1985)	Probit	2	9	No	No	No
Main & Shelly (1990)	Probit	2	14	No	No	Yes
Whitfield & Bourlakis (1991)	Probit	2	10	Yes	Yes	No
Main (1991)	Probit	2	14	No	No	No
O'Higgins (1994)	Switching Regression	2	22	Yes	No	Yes
Dolton et al. (1994c)	Probit	2	10-14	Yes	Yes	Yes
Payne (1995)	Multinomial Logit	5	16	Yes	No	Yes

[1.] *Not including regression constant*

4.5: Training Provision in Australia

In the last decade Australia has made significant advances in improving vocational education and training standards. This process has included significant reform of its vocational education and training governance structure and the introduction of a nationally consistent qualification's framework (Stern, et al., 1996, pp. 28-29). These reforms bring Australia's school-to-work transition arrangements more in line, at least in scale, with arrangements in other OECD countries (Sweet, 1999, pp. 71-72). Sweet notes that a number of significant historical influences preceded this push for reform and impact on the arrangements now in place. First, Australian states and territories only agreed to join together as a nation state at the beginning of the twentieth century resulting in the formation of the Commonwealth (Federal) government responsible for economic policy. The six States and two Territories had by this stage however, already developed a degree of independence in the way they governed their affairs including education and training delivery. That independence has subsequently shaped and, at times, hampered the harmonisation of national education and training policies. Second, the pattern of early settlement in Australia led to the importing of some class conflict and sectarianism which has helped shape its educational and training institutions. For example, Australia has one of the world's largest private school systems much of it sectarian in origin, catering for one in three of all students in the final two years of secondary school. Similarly, its initial vocational preparation system was constrained by the tradition of viewing education and vocational training as separate activities for young people. Such traditions led to separate bodies within each State and Territory being responsible for the setting of standards and certification of outcomes in secondary schools, technical colleges and training. These developments led Sweet to observe that

> "The combination of this institutional polarization within each state with a division of powers between the Commonwealth and state governments has given an unusual degree of complexity to the governance of post-compulsory education and training in Australia" (ibid. p.71).

Third, Australia adopted a centralised wage bargaining system in which craft based unions played a pivotal role underpinning institutional arrangements such as minimum wages, national awards and apprenticeship arrangements. This included a system of junior wages providing up to a five-year cushion for the inexperience of school-leavers to allow the acquisition of knowledge and skills expected of an adult worker through work experience. Fourth, Australia's use of a large migration programme to fuel post-World War II industrial development resulted in a heavy reliance on imported labour skills rather than on domestic education and training (op. cit., p. 69-72).[95] Sweet argues that these features of Australian heritage, combined with a resource-based rather than manufacturing-based economy, "...helped reduce

---

[95] Sweet (1999) cites a 1978 OECD report *Intergovernmental Meeting on Vocational Education and Training: National Report on Australia,* which estimates that migration was responsible for 40 percent of the annual supply of skilled trade-persons as recently as the mid 1960s.

the urgency of considering the development of human capital as part of national economic policy" (p. 71).

This meant that, until the early 1970s, formal education and training ceased for the majority of young Australians at the minimum school leaving age of 15 with fewer than 20 percent completing the last two years of secondary school. Approximately 10 percent of each cohort entered apprenticeship training but the labour market was able to provide most of the remainder (including immigrants) with jobs that required no qualifications beyond the completion of two or three years of secondary education. The school-to-work transition was supported by the Commonwealth Employment Service (CES), which was established in 1945 and designed to assist job search and placement. The junior wage system, which began at a low level for 15 year olds but progressed in annual increments until reaching adult levels at around 20 years of age, allowed necessary knowledge and skills to be acquired in the workplace (Sweet, 1999, p. 72). Informal on-the-job training thus represents Australia's largest training resource while the provision of formal training has historically relied, almost exclusively, upon employer-led initiatives, particularly apprenticeships which in 1990 provided training for approximately 23 percent of all employed 15 to 19-year olds (Baker, 1994, p. 71).[96]

The Australian apprenticeship system of training naturally grew out of Australia's British heritage together with the rapidly increasing demand for skilled trades-persons in newly founded industry and large-scale investment in infrastructure. In the second half of the nineteenth century, Mechanics' Institutes, Schools of Mines and Technical and Working Men's Colleges were formally established to develop the skills for the traditional craft occupations in the metal, electrical, mechanical and construction trades (ANTA, 1999). The period of indenture is typically 4 years and wages, protected by the award system, rise from 40 percent of the adult wage in the first year, to 90 percent in the final year. Apprenticeships are, however, predominately the preserve of male youth and even significant recent growth in

---

[96] *Formal* is usually taken to mean training that leads to a qualification; as distinct from *informal* that does not.

service trades such as cooking, hairdressing and jewellery making has brought only minor increases in female participation (Sweet, 1999, p. 93). While the Australian apprenticeship system maintains many of the features of the traditional UK apprenticeship it has managed to adapt to change more readily than its British counterpart (Gospel, 1994). Reforms introduced in 1998 have, however, substantially changed the nature of apprenticeship in Australia, in terms of tenure and delivery, in an attempt to meet the challenges of the twenty first century (see below). The numbers on apprenticeships between 1985 and 1997 are shown in figure 4.3 (together with commencements for both apprenticeships and traineeships). Although commencements peaked in 1988/89 and 1989/90, apprenticeship training in many respects, continues to be the backbone of formal training arrangements in Australia, the scale of which is only surpassed by the German speaking countries and Denmark where such schemes have a broader occupational coverage (Baker, 1994, Sweet, 1999, see also take up of 'New Apprenticeships' in table 4.5 below).

The Commonwealth Government has actively assisted the apprenticeship system through a series of initiatives dating from the early 1970s including financial rebates and grants to employers. These were designed to encourage release for off-the-job training, use of spare training capacity, greater occupational coverage and, more recently, the taking on of young people classified as disadvantaged in the labour market.[97] Additionally, apprentices are eligible for living away from home allowances to encourage greater mobility and funding for approved off-the-job-training courses (ABS, 1985, p. 155; ABS, 1998, p. 220). Group apprenticeship schemes, where apprentices are leased out to several employers during the indenture period, also received additional support. Preparatory courses were also supported through the payment of a pre-apprenticeship allowance when other Commonwealth assistance is not available at a higher rate. The government also sought to address skill shortages identified in a limited range of occupations through the establishment of a further entry level initiative, *Skills in Demand* (later renamed National Skills

[97] Identified by the CES as including Aboriginal and Torres Strait Islander peoples, people with disabilities, youth and sole parents. The Australian National Training Authority also includes women, young offenders and residents of rural and remote communities (ANTA, 1997, p.12).

Shortage), which provides the costs of establishing and running training courses and providing allowances for unemployed trainees.

Like the UK, Australia experienced significant structural change in the 1960s and 1970s and the onset of recession following the 1970s oil shocks, but it also faced significant constraints in traditional export markets as a result of the UK joining the Economic Community. Youth unemployment soared, from less than 3 percent in 1974 to over 10 percent in 1979, and it has subsequently been a constant backdrop to post-compulsory training and education policies (Sweet, 1999, p. 73). The traditional apprenticeship system, confined to a narrow occupational base, struggled to deliver the breadth of skills needed in a more open and competitive environment. Moreover, apprenticeships were inflexible and entry was highly sensitive to the state of the economy (Baker, 1994, pp. 71-72). Employers had to respond by training without the apprenticeship system, particularly in the rapidly developing service sector. Problems in the training system were, however, exacerbated by the lack of a cohesive nationally recognised framework. Every state and territory had its own system, with different courses, qualifications and rules. This made it difficult for people to move inter-state, as often their qualifications/skills were not recognised (Hall, 1995; ANTA, 1999).

A number of training initiatives were introduced in the 1970s but these were predominately employer led, some were formalised with certificates, but it was not until the 1980s that formal courses, qualifications and rules were established with State and Territory government backing.[98] In 1985 the Labor Commonwealth Government, in an attempt to introduce greater flexibility and broader coverage into contractual working arrangements, introduced the Australian Traineeship System (ATS). Although take-up was initially slow, as illustrated in figure 4.3, *traineeships* were designed to complement the apprenticeship programme and share many features with apprenticeships. Primarily they continue to rely on industrially negotiated contracts of employment and training, however they only last for one year

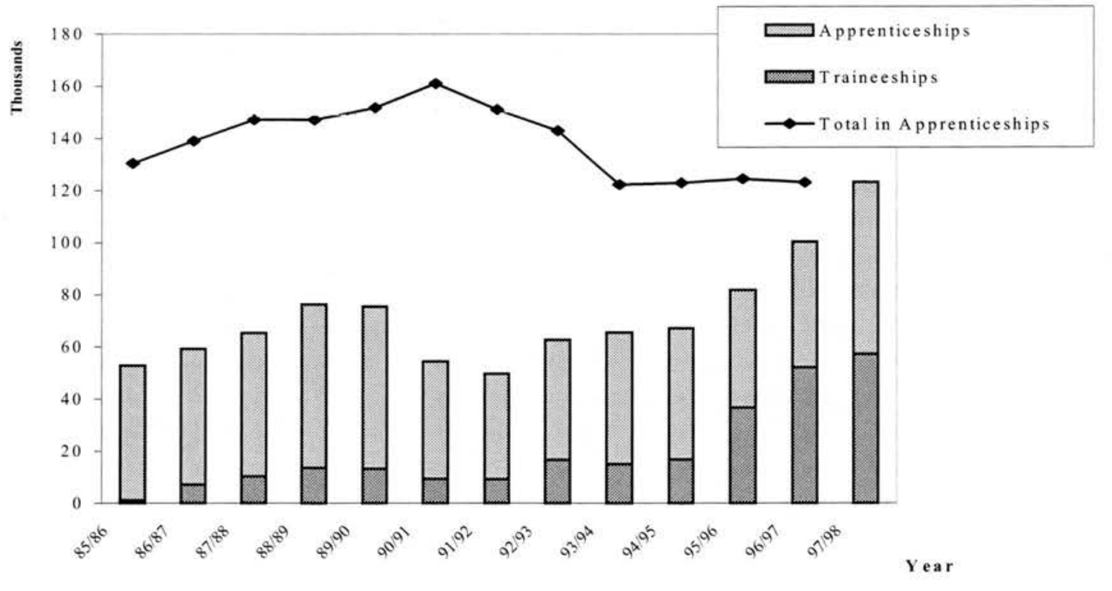
[98] A notable exception was the Trade Union Training Authority established in 1975 for the provision co-ordination, promotion and evaluation of trade union training in Australia (ABS, 1985, p. 158).

(although some later contracts were extended to two years) and the employer is not obliged to bear the cost of time spent in off-the-job training. Preference was given to occupations and industries other than traditional trades such as those in office-based and retail industries. In recent years, however, they have also become established in industries that traditionally used apprenticeships. The off-the-job training component, usually in the Technical and Further Education (TAFE) sector, typically comprises the equivalent of two days a week over the year and was funded by the Commonwealth government. A training fee of $2,000 was paid to employers to offset the cost of on-the-job training and additional incentives were available for taking on disadvantaged workers and successful completion (ABS, 1995, p. 204). These initiatives took place in a period of significant economic reform leading to a more open economy including wide-scale deregulation, award restructuring, privatisation, significant tariff reduction and the introduction of a floating exchange rate.

Pressure for increased government intervention in the late 1980s was motivated by an emerging consensus among policy advisors, industry representatives and trade union leaders that Australian industry suffered a competitive disadvantage because of under-investment in training (Baker, 1994; Hall, 1995; ANTA, 1999). A number of factors were identified as contributing to the inadequate training levels including the nature of award structures, high labour turnover, poaching, poor management attitudes and a reliance on immigration (Chapman and Stemp, 1992, p. 355). Significantly, awards, which were organised along occupational lines, established minimum wages and working conditions still cover between 80 and 85 percent of employees in Australia.

Baker (1994, p. 74) argues that the award system inhibits the process of training based skill acquisition due to the rigidly defined occupational classifications, high minimum wages and well-defined promotion processes. These rigidities, along with well-defined wage relativity's between different categories of workers, combined to limit the range of tasks the typical trainee had to perform, constrain worker wage profiles and discourage investment in training by firms. Moreover, the state of the youth labour market continued to decline in the 1980s with approximately 20 percent of teenagers and 10 percent of young adults experiencing unemployment (Miller and Volker, 1987, p. 203). This resulted in further significant training reform taking place from the late 1980s in an attempt to establish a 'training culture' in Australian industry under the governance of the Department of Education, Employment, Training and Youth Affairs (DETYA).

Figure 4.3: Commencements in traineeships and apprenticeships, Australia: 1985-98



Source: 1985-1995: ANTA figures see Ainley et al. (1997)
1996-1998: ANTA figures see NCVER. (1999)

Note: Year ending June 1998 breakdown is approximate only (my calculations) as the distinction between apprenticeships and traineeships was formally abolished with the introduction of New Apprenticeships on the 1st of January 1998. See also table 4.5 below.

For example, in 1989 the Commonwealth Government introduced the *Industry Training Levy* (ITL), under the *Training Guarantee Scheme*. The ITL required firms with a minimum payroll to spend 1 percent of payroll on approved training activities or be taxed an equivalent amount.[99] If the firm did not undertake the required

[99] The levy did not, however, come into force until July 1990 and applied only to firms with a payroll exceeding the threshold of $200,000 (indexed to average weekly earnings). Provision existed to exempt firms in the building and construction industry as well as shearing and related occupations (ABS, 1995, *Yearbook Australia,* p. 205).

training expenditure, subsequently increased to 1.5 per cent in July 1992, then it was required to forfeit it as a levy (collected by the Australian Tax Office) to the training authorities (DETYA, *Annual Report*, 1996). It was thus similar, in many respects, to the UK's ITB levy scheme, which was abolished in 1979. The underlying motivation for the ITL was the widely held perception that the private sector was inadequately providing for skill formation (Baker, pp. 81-84; Chapman and Stemp, 1992, pp. 354-355). An evaluation by the DETYA claimed that the scheme was particularly effective in maintaining the level of training activity during the 1990-93 recession, and in creating awareness of training as a strategic business issue. The study revealed that more than half the eligible employers increased their training expenditure as a result of it. Nevertheless, the DETYA determined that the ITL was not successful in establishing a training culture or helping disadvantaged groups and it was subsequently suspended for two years from 1 July 1994 before being abolished by legislation, which was passed in 1996.

The National Training Board was established in 1990 and, in 1992, the Commonwealth, State and Territory governments agreed to establish the Australian National Training Authority (ANTA), charged with developing and overseeing national priorities and goals in training. A Board of prominent industry representatives steers the ANTA to ensure a consistency between national training strategy and the needs of industry (Hall, 1995, p. 87). The ANTA is also responsible for advice on the distribution of Commonwealth funding between States and Territories, continuing development of the national training framework, policy review, evaluation, and research on national priorities, as well as the implementation of any major national (training) policies agreed by Ministers (ANTA, 1999). The National Training Board assists in this process through the governance of a system of recognised training providers, linked in with the endorsement of competency standards across industries or agreements determined by an industrial tribunal.

ANTA facilitated the development of an Australian Vocational Certificate Training Scheme (AVCTS) in 1992 to promote a more unified entry-level training system (Ainley, Malley and Lamb, 1997). Significantly, the AVCTS is based on competency based training with nationally determined standards in consultation with industry bodies.[100] One of the key objectives of the AVCTS is to formalise previous informal training and increase available training credentials (Baker, 1994, p. 72). It has introduced a graded certification system that covers different forms of vocational education including training delivered on-the-job and through the TAFE system. The government funded TAFE remains the primary provider of formal (off the job) training and certification delivering an estimated 85 per cent of Australian training, through 84 institutes in 1994/95. The remaining 15 per cent of training in this period was delivered by approximately 2,500 registered private providers and an unknown number of unregistered ones (ANTA, 1999).[101] However, while the majority of TAFE entrants study part-time and are employed, the diversity of courses offered by the institutes' mean that not all courses can be regarded as job-related in their training content (Baker, 1994, p. 72). The ANTA initiatives helped to significantly boost numbers on Traineeships in 1993 and 1994 (see figure 3.3) and its establishment has led to greater co-ordination between State and Territory training authorities (Ainley, et al. 1997, p. 29). These efforts, together with initiatives in the educational sector, have culminated in the inclusion of vocational training assessment within the Australian Qualifications Framework (AQF). In other words, the establishment of the AVCTS has helped achieve a national focus in the setting of training standards and accreditation of courses. Moreover, ANTA also assists employer provided training through the provision of assessment guidelines and registration of training providers.

---

[100] Baker (1994) defines competency-based training as "...training, which is designed to focus on *outcomes* such as the attainment and demonstration of knowledge, skills and applications, as opposed to training *inputs*, such as time served" (p. 72). Significantly, many traditional apprenticeships were based on time served, and thus much of the training content rested on the concept of 'learning by doing'. The official definition of competency-based training is, however, more focused. The National Centre for Vocational Education Research (NCVER) describes it as training "...aimed at the acquisition of knowledge and skills, and their application, to meet industry-specific standards rather than individual's needs" (NCVER, 2000, p. 105). See also Stern et al. (1996, pp. 16-22).

[101] The NCVER (2000) report on training statistics for 1999 listed 85 TAFE Institutes (over some 1,132 locations), 1075 community education providers and 2465 other (predominately private) providers.

The Australian government continues to concentrate its youth training support on employer-led, contractually based initiatives which have their origins in the traditional apprenticeship system imported by European settlers more approximately two centuries earlier.[102] The latest adaptation, the *New Apprenticeship* system,[103] is designed to "...ensure that Australian enterprises have access to a pool of skills of world-class quality" (DETYA, *Annual Report*, 1999, *Sub-programme 3.1*). With its introduction on the 1st of January 1998 the distinction between apprenticeships and traineeships was formally abolished introducing greater flexibility in terms of tenure and qualifications. The system also seeks to increase the occupational coverage and flexibility of apprenticeships and encourage new opportunities in non-metropolitan Australia. It provides an additional $1,000 incentive to employers for each new Apprentice achieving AQF Certificate III or IV in trades and occupations identified as having a skills shortage.[104] Incentives for group schemes have also been substantially increased to encourage mobility of apprentices and greater participation of small and medium sized enterprises.

Statistics from the National Centre for Vocational Education Research (NCVER) suggest that the initial impact of the New Apprenticeships appears to have been positive with an estimated increase in commencements of 35.4 percent being registered in the first full year of operation.[105] Significantly women made up 41.3

---

[102] That is, support of Apprenticeships, Traineeships and Special Trade Training (targeted assistance to establish apprenticeships in non-traditional occupations) - classified by the Australian Bureau of Statistics as 'Entry Level Training'.

[103] Formally titled the Modern Apprenticeship and Training System (MAATS). Schofield (1999b, p. 1) notes that the national decision to encompass both apprenticeships and Traineeships by the single term of *New Apprenticeships* has led to some confusion and that, in practice, the separate terms continue to be used by many governments, industry and the wider community. Moreover, statistical analysis of apprenticeships and Traineeships as distinct systems will become more difficult as previously separate statistics on the two schemes are progressively replaced by composite statistics on the New Apprenticeships.

[104] The Australian Education Council has set targets for education and training in terms of AQF levels. Specifying that by 2001 60 per cent of 22 year olds should either be participating in education and training programmes which lead to level III awards (e.g. a trade certificate), have attained level III (or above) qualifications or be participating in (or have completed) higher education studies such as degrees and diplomas.

[105] Over the combined annualised figures for Apprenticeships and Traineeships in 1997. The NCVER has collected and collated a range of labour market statistics on behalf of ANTA since 1992.

percent of the 1998 trainees commencements, reflecting the broader appeal and wider occupational coverage of the new scheme (NCVER, 2000). Moreover, the uptake of older workers (20 years and over) has been enhanced through the widening of Commonwealth assistance to include existing workers whereas it was previously only available to young people entering the labour-force (see Schofield, 1999a). DETYA figures on the initial uptake of New Apprenticeships are shown in table 4.5 by AQF level. In a parallel development youth are being encouraged to undertake vocational education and training while still at school through the *Australian Student Traineeship* scheme and *Job Pathways* programme. Both of these initiatives have been set up with the explicit objective of assisting students in making 'sustainable school-to-work transitions' (DETYA, *Annual Report*, 1999, *Programme 3*).

Table 4.5: Australian New Apprenticeship commencements by AQF level, 1997-1999

AQF Level	*1997*		*1998*		*1999*	
	number	%	number	%	number	%
Diploma or higher	116	*0.1*	207	*0.1*	296	*0.1*
Certificate IV	1,733	*0.9*	6,706	*3.1*	7,588	*2.7*
Certificate III	131,552	*70.5*	155,991	*71.2*	194,119	*74.7*
Certificate II	41,318	*22.1*	49,715	*22.7*	54,176	*20.8*
Certificate I	573	*0.3*	281	*0.1*	152	*0.1*
Not Known	11,426	*6.1*	6,085	*2.8*	3,550	*1.4*
Total	***186,718***		***218,985***		***259,881***	

Source: DETYA Annual Report, 2000, pp. 42-43 and table 10.

Although the emphasis of Australian training has been on work-based training initiatives, a variety of small-scale labour market programmes closely approximating the British model have also been used in Australia to buttress the employer-led initiatives as unemployment rates rose in the 1970s. The majority of these initiatives were designed to improve the employment prospects of the long-term unemployed and other disadvantaged job seekers (see footnote 97 for the formal CES definition of disadvantage) and thereby both the equity and efficiency of the labour market. Although few of the programmes have been designed exclusively for the young, they tend to make up a disproportionate number of entrants relative to their share of registered unemployment (Sweet, 1999). A selection of different labour market programmes established since the late 1970s that young Australians could benefit from are shown in table 4.6 (the current programmes, or variants, are shown in *Italics*).

Programme assistance ranges in function from subsidised employment to the purchase of formal training places in TAFE colleges or other registered training providers. While the Commonwealth government funded the bulk of labour market programmes, State and Territory governments retained responsibility for accreditation and certification as well as making some additional training places available. The State and Territory roles were however conducted with little co-ordination with the Commonwealth programmes. Moreover, unlike apprenticeships and traineeships, labour market programmes fall outside the negotiation brief of industry partners leading to award settlement, either in relation to their content or the qualifications to which they lead (Sweet, 1999, 110-111). The ANTA is charged with resolving these ambiguities as Australia moves toward a single national framework for vocational education and training (Stern, et al. 1996).

Table 4.6: Australian Commonwealth Government labour market programmes

Programme Type	*Eligibility*	*Programme description*
Employer Incentives		
Jobstart	Disadvantaged jobseeker or long term unemployed	wage subsidy scheme subsidy varies with age, unemployment duration & degree of disadvantage duration of 26 weeks
Training for Employment		
Standard SYETP [a]	15-24 year olds, unemployed and out of education for 4 of the previous 12 months	subsidised work experience subsidy varies with age duration of 17 weeks
Extended SYETP [a]	18-24 year olds, unemployed and out of education for 8 of the previous 12 months	subsidised work experience subsidy varies with age duration of 17 weeks + 17 more at reduced subsidy
Skillshare	Disadvantaged job seekers and long-term unemployed	grants to non-profit community groups structured skills training short term
Jobtrain	Disadvantaged job seekers and long-term unemployed	short-term vocational training duration of 6-8 weeks
Jobskills	Long term unemployed (12 months or more)	work experience and training duration of 26 weeks
New Apprenticeship Access Programme	Disadvantaged job seekers and long-term unemployed	pre-vocational training duration flexible
Employment Creation		
Community Employment Programme (1983)	Disadvantaged job seekers and long-term unemployed	70-80% funding of labour intensive community project costs where additional employment is created
Community Youth Support Scheme	Unemployed youth (registered)	community or voluntary sector based work experience and training funding of labour intensive community projects
LEAP [b] (Now Green Corps)	Unemployed 15-20 year olds	work experience and training in landcare, cultural heritage & conservation activities duration flexible
Job Seeker Preparation		
Workplace English Language and Literacy	Individuals referred by CES or other agencies	training in language proficiency for occupational purposes
Job Pathway Program	School leavers	career guidance and job placement and mentoring support duration 6 months
Special Training		
Aboriginal Employment	Aboriginal and Torres Straits Islanders	work experience and training institutional training project funding duration flexible

a. Special Youth Employment Training Program
b. Landcare and Environment Action Program

Significantly, the most recent labour programme developments indicate a policy shift toward working more closely with the employer-led training system, specifically supporting programmes that "...contribute to the development, implementation and take-up of New Apprenticeships" (DETYA, *Annual Report*, 1999, *Programme 3*, see also Ainley et al. 1997, pp. 34-35). For example, the Jobs Pathway programme has the flexibility to fund a disadvantaged job seeker through a New Apprenticeship. Moreover, recent DETYA *Annual Reports* clearly place emphasis on this link by reporting the performance of Jobs Pathway in terms of placements in apprenticeships and traineeships.

The CES, like its British equivalent, also provides assistance in job search activity, such as mobility grants, interpretation assistance, counselling services, and job placement. It is also responsible for a number of more general training programs, such as the *Workplace English Language and Literacy Programme*. This programme was introduced in 1991 and assists workers in gaining the language and literacy skills sufficient "...to enable them to meet the demands of their current employment and their future employment and training needs" (ABS, *Yearbook*, 1995, p. 206). It reflects the continued reliance of the Australian labour market on immigrant labour and the desire of the ANTA to equip all young people with the necessary tools to enable participation in apprenticeships and traineeships. Basic training in information technology is also provided for job seekers where needed (DETYA, *Annual Report*, 1996). In a related move the CES recently placed all its job vacancies on their internet site, accessible via terminals in all CES centres, which has the incidental benefit of introducing computers to some job-seekers and enhancing the computer skills of others

In summary, the Australian youth training system has experienced considerable reform since the late 1980s, which has witnessed emergence of a cohesive national vocational training framework, in terms of delivery and governance. The Australian government has chosen not to depart from the traditional apprenticeship system to the extent that has occurred in the UK. Rather it continues to adapt and modify its apprenticeship system to meet the training needs and objectives of industry. Hence, the vast majority of training activity in Australia remains employer-led, and hence takes place in the private sector, but as Baker (1994, p. 71) notes it is conceptually difficult to disentangle the role of industry from the roles played by government and educational institutions. This overlap will continue to grow as the level of integration between Australia's small-scale labour market programmes and its employer-led training system increases under the stewardship of the ANTA. The main purpose of intervention in the form of the complementary labour market programmes is to increase equity for disadvantaged groups and the long-term unemployed. This encompasses support to enable recent immigrants to participate in apprenticeships. Apprenticeships and traineeships also receive direct government support on a number of levels ranging from subsidies to the provision of off-the-job training. Moreover, institutional structures provide further support in sustaining the apprenticeship system, particularly the award system.

4.6: Evaluation in Australia

The vast majority of evaluation activity on government supported training in Australia continues to rely upon aggregate measures. The DETYA reports placement rates for different programmes in its *Annual Report* and commission occasional evaluations based on activity measures following programme participation. For example, Byrne (1993) evaluated the performance of Jobstart who completed their wage subsidy period in 1992 and were surveyed again 6 months to estimate the employment gain relative to a comparison group. Employment rates of participants were twice that of non-participants (60 percent compared with 30 percent) but the study did not control for sample selection bias. Another department report (DETYA, 1991) followed Australian Traineeship Scheme (ATS) participants for three years after completion of training, reported that 87 percent of respondents had work in each survey year. Significantly more than 50 percent of the respondents remained with the training employers but again no attempt was made to follow or construct comparison groups.

In a series of recent articles Schofield (1999a, 1999b, 2000) reviewed the performance of ATS in a number of different States.[106] Focusing on the degree of satisfaction amongst employers and trainees with various aspects of Traineeships, the reports found a generally high overall level of satisfaction with the scheme (see for example, Schofield, 1999, pp. 3-4). Schofield identified some concerns associated with the quality of training and assessment, including evaluation of the scheme as a whole. However, little evidence was found to support concerns that training quality had been traded off for Traineeship quantity (op cit., 2000, p. iv). The major cause of concern in the Traineeships was the relatively high non-completion rates, which for example, averaged approximately 30 percent in Tasmania between 1990 and 1998 (ibid. p.13).

Given that bulk of training provision in the Australian youth labour market is considered to be 'employer-led' it is not surprising that the majority of econometric evaluation activity has concentrated on returns to private sector training. The notion of an employer-led system, however, raises an interesting conceptual issue, previously considered by Dolton, et al. (1994a, pp. 204-205). For instance, it does not make it clear what the alternative to an employer-led system is. If we consider the Australian system to be employer led, does this for example imply that the UK system is government led? Moreover, it is not clear that a specific agent can be identified as the 'leader' of a particular training spell, particularly if the training is provided within some sort of market structure, which, by definition, should ensure that all parties are equitably compensated. Dolton, et al. suggest one way to overcome this difficulty is to divide the training spells being analysed up into a number of subsets dependent upon the nature and financing of the training (ibid. p. 205). This approach was subsequently applied to evaluations of the United Kingdom's YTS by these authors (Dolton et al., 1994b) and others (Payne, 1995; Green et al., 1996).

The main reason that the Australian system attracts the employer-led tag is, however, that the Traineeship system (like apprenticeships) is contractually based. The

[106] The relevant states are Queensland, Tasmania and Victoria.

contract may oblige the employer to employ the trainee for a set period after completion. This will impact on short-term evaluations artificially raising employment outcomes and perhaps lowering wage outcomes. Equally, however, implicit contacts may perform in the same way in the UK and a number of studies have emphasised the necessity to control for the case when a trainee remains with the same employer who provides them with the training (see for example, Main and Shelly, 1991).

When the role of various institutions is considered the difference between government and private provision becomes even more difficult to distinguish and it is thus necessary to make a variety of assumptions. For example, Miller (1990, p. 15) observes that it is difficult to determine, *a priori*, if wages set under the [Australian] award system obstruct rather than mimic market forces. These issues will be discussed further in later chapters but it is important to emphasise that the key feature shared by Australian Traineeships and Youth training in the UK is that both systems are primarily work based. It is this feature that provides a basis for comparison of youth training in these countries. In the remainder of this section a selection of recent Australian econometric evaluations are reviewed preceded by a brief description of available data sources in Australia.

4.6.1: Data Sources in Australia

Australia has a well-established longitudinal survey programme beginning with the surveying of a cohort of long-term unemployed youth registered with the CES in 1984. This was followed by an interview-based survey, the Australian Longitudinal Survey (ALS), of a representative sample of the population aged between 16 and 25 years conducted annually until 1991. This was followed by the Australian Youth Survey (AYS), which began in 1989 with a nationally representative sample of individuals aged between 16 and 19. This group has been re-interviewed each year since with a new group of 16 year-olds added each year from 1990 to 1994. The aim of the AYS was to enable the analysis of the rapidly changing situations in the youth labour market with respect to increased school staying on rates and the school-to-work transition. Labour market topics include detailed work history, job training, job search behaviour, and experience with the CES. Other topics related to the main labour market theme include secondary schooling, further and higher education, career advice, qualifications, school-to-work transition, health, housing and financial conditions. The survey also includes basic demographic variables such as age, gender, marital status, racial origin, parental education and occupation and area of residence.[107]

The AYS provides a comprehensive database on the labour market behaviour of a representative sample of young Australians. However, with respect to the evaluation of training schemes, Wooden (1995, p. 118) makes the observation that "...training is only of minor significance within the overall context of the survey".[108] The only measures of training are simple binary variables indicating participation in different types of training during the year prior to the interview. The data thus provides little guide to the extent of differences in the quality and relevance of the training received. This weakness is partly abated by the distinction between public and private sector training, comprehensive coverage of apprenticeships and job search behaviour, together with detailed coverage of further and higher education. The role of career guidance, both at school and by the CES, is given considerable weighting and data on wages and hours worked is of good quality.

Overall, many of the variables observed in the UK datasets can be identified in the AYS and sufficient degrees of freedom maintained in a comparative study. The dataset lends itself to examining training as only one of several alternative school-to-work routes, as employed in the multivariate analyses of Dolton, et al. (1994b); Green, et al. (1994); and Payne (1995). Multivariate analysis is further enhanced by the length of coverage of the AYS allowing the training route to be compared with transition routes of longer duration, such as higher education and apprenticeships. Given the detailed nature of the calendars and range of questions addressing job search, the dataset also lends itself to duration analysis as employed by Dolton, et al. (1994c). Moreover, Australian institutional characteristics closely approximate those

[107] The AYS is described in more detail in the appendix to this thesis.

[108] Partially due to this Wooden (1995) further notes most Australian research to date has used alternate data sets.

in the United Kingdom with a number of broadly similar youth labour market reforms being adopted although the timing has tended to differ. Indeed, the richness of the Australian dataset potentially allows a wider range of possible specifications however, in practice, these are limited by a relative paucity of observations detailing the receipt of government assisted training. The remainder of the thesis hence concentrates on the SYPS.

4.6.2: The Estimated Impact of Training

As observed previously, econometric estimates of training impact are largely concentrated on the private sector. For example, Miller (1990) assessed the impact of different types of training on earnings of young workers using the 1985, 1986 and 1987 sweeps of the ALS. Miller utilised three methods, OLS, Heckman's two-step correction model and instrumental variables to estimate the impact of training on young people's earnings. When applying the Heckman model to allow for sample selection, he found that the coefficient of the correction term (Heckman's lambda) was insignificant in both the participant and non-participant equations and the resultant estimates were thus similar to those produced using OLS.[109] Similarly, he finds that the instrumental variable approach yielded results closely approximating the OLS estimates and hence his discussion concentrates on the latter. Miller fails to detect a significant wage premium associated with training. He attributes this to: the lack of precision in the binary training index, the distribution of respondents across the training and post-training states, and the short average employment tenure within the sample (two and one half years) (Miller, 1990, pp. 16-17). He tests a number of alternative specifications however, consistent with a number of UK studies, Miller finds that job tenure (experience) and educational attainment have greater explanatory power in determining outcomes. He notes, however, that the level of training is reasonably small, perhaps amounting to around two hours per week for training recipients (ibid. p. 20).

---

[109] Returns to higher educational attainment and general labour market experience were however higher in the participant equation. This is consistent with Miller's examination of the incidence of training across the sample, which showed that training and formal education are complementary (Miller, 1990, pp. 7-12).

Tan, Chapman, Peterson and Booth (1991) also draw on ALS data to compare the impact of training in Australia with training outcomes in the UK and US in the late 1980s. After establishing a broadly consistent definition of schooling completion they treat all reported training events (including school-based training) as episodes of post-school training using the data to describe and compare the incidence, timing and impact of training. In each of the countries schooling attainment was confirmed as the key predictor of post-school training. The study found that American youth were less likely to participate in formal training than youth in the UK or Australia with just 12 percent reporting having received training compared with 30-40 percent in the other countries (50-60 percent if apprentices) (Tan et al. 1991, pp. 14-16). They are, however more likely to receive on-going training in the workplace, whereas British and Australian youth training tends to be concentrated in the initial years of employment (ibid. pp. 7-9). In terms of the impact on earnings, training appears to yield higher marginal returns to labour market entrants with low initial stocks of human capital in each country. The returns are much higher in the US, ranging from 10 to 18 percent for off-the-job and company-based training in the US compared with 4 to 10 percent in Britain and Australia.[110] Consistent with Miller (1990), work experience and educational attainment are identified as the major determinants on future earnings in the US and the UK as well as Australia. The impact of training on employment stability (the probability of unemployment) is found to be positive in all three countries, regardless of the type of training.

Given the emphasis on programmes for job seekers considered to be at risk or disadvantaged a number of evaluations have specifically examined the effect of being an Indigenous Australian on employment. For example, Miller (1989) used ALS data to analyse the unemployment differential between indigenous and non-indigenous youth in 1985. Utilising a Probit equation to control for individual and background characteristics, such as education, age and location, he estimated that Aboriginal youth were two and half times more likely to be unemployed than other

[110] The categories of training differed between the individual countries but can be broadly defined as company-based, off the job, or at school. The latter category was included for both the US and the UK but the study did not identify any measurable effect on wages. The authors also note that the inclusion of shorter training spells in the Australian data may bias the relevant estimates downward.

groups.[111] Using data from the 1986 and 1991 censuses, Daly (1993) found that this pronounced differential persisted even after accounting for the lower than average educational attainment of Aboriginal Youth. After controlling for education, geographical location and various background features she estimated that a 'change in race' from non-Aborigine to Aborigine is associated with a reduction in the probability of employment by 23 percentage points for males and 21 percentage points for females (ibid. pp. 145-146). Daly did, however, find evidence suggesting that the Communities Employment Programme may have reduced the previously observed differential in employment probability between rural and urban Aborigines.[112]

The most recent initiative is the report released by the DETYA Task Force on School-to-Work Transitions for Indigenous Australians. The report used 1991 and 1996 census data and concentrated on aggregate measures but provides a recent profile of patterns in the school-to-work transition of Indigenous Australian youth relative to their non-Indigenous peers.[113] They are under-represented, in terms of participation levels, in post-compulsory schooling, further and higher education. Indigenous Australians are, however, more likely to opt for the vocational education and training route leading to an over-representation of aborigines in Traineeships and other government assisted schemes.

At nearly every stage, Indigenous youth experience substantial disadvantage in terms of their participation in education, their attainment of educational qualifications, and their participation in the labour market. Moreover, in most instances, this disadvantage is substantial. The report identified variation in the extent of disadvantage dependent on gender and location but only rarely does it disappear.

---

[111] The relative Aboriginal and non-Aboriginal estimates were 26.11 percent compared with 10.54 percent for males and 24.85 percent compared with 11.05 percent for females (Miller, 1989, pp. 47-48).

[112] 1986 census data revealed that only 24 percent of indigenous Australians lived in major urban areas compared with 64 percent of non-indigenous Australians (Daly, 1993, p. 135).

[113] The report notes that the comparisons between the 1991 and 1996 figures are limited by the short time period, a possible increase in the extent to which persons were willing to identify themselves as Indigenous and the reliance of placement rates alone. Moreover, the report observes that the results are often 'equivocal in any case' (DETYA, 1999, p. 36).

While higher educational attainments are associated with improved labour market outcomes the report disturbingly finds that disadvantage persists between Indigenous and non-Indigenous young people even where they have achieved the same level of educational attainment and live in the same location. This suggests that other factors, not addressed by the report, may be responsible for persisting disadvantages including, for example, poorer levels of health and discrimination. That is:

> "[such] …results indicate that improved educational outcomes alone are not sufficient to remove the inequalities experienced by Indigenous Australians" (DETYA, 1999, p. 36)

The lower levels of educational and labour market participation observed for Indigenous youth continue despite government and institutional policies and programmes, such as special entry programs in TAFE or universities and the Communities Employment Programme, all of which are designed to improve outcomes. The report notes, however, that were it not for these programmes, it would be reasonable to expect the level of disadvantage experienced by Australia's Indigenous youth to be substantially greater than it is.

4.7: Chapter Summary

In the post-war period youth training in the United Kingdom and Australia has undergone considerable institutional reform typified by increased government provision and funding in both countries. This chapter has traced these reforms in the face of the considerable structural, technological and demographic changes that have impacted on both the demand for and supply of young workers. In general, the aim of the reform has been to encourage young people to stay in secondary school for longer or to undertake further education or vocational training leading to a nationally recognised qualification. To a great extent, this has resulted in a blurring of the boundary between education and training, with some vocational courses being available within secondary schools. Traditional work-based routes, such as apprenticeships, have been extensively modernised to place greater emphasis on competency-based training and allow greater flexibility. As such the school-to-work transition has become more difficult and complex. Although training systems in the UK and Australia are similar in many aspects two key differences are however readily apparent. First, Australia continues to rely on employer-led initiatives more than in the UK, in the sense that traineeships are contractually based. Second, Australia has developed a large number of smaller scale government sponsored youth training programmes that are targeted to specific client groups, predominately indigenous Australians.

Evaluation studies of British youth training programmes have employed a range of economic and statistical specifications. The resultant estimates on the impact of training programmes have however tended to vary widely, particularly when measured in terms of the probability of gaining employment. The estimates are also subject to a number of biases undermining their explanatory power and stability. Different sources of bias have been identified, ranging from those inherent in the datasets, to those in the choice of model variables and structure. However, attempts to correct for these have tended to be piecemeal and varied, in addition to being applied to different surveys and survey sweeps. This has made it difficult to draw any firm conclusions regarding the practical importance or relevance of these attempts to correct for the various sources of bias. Subsequent research should be aware of and attempt to take into account a wider range of potential bias. Given the wide range of potential specifications, a sensitivity analysis, applying an array of feasible specifications to a common database with appropriate testing, would be of considerable benefit in helping other researchers determine which techniques and specifications are more relevant in different circumstances. This is the approach adopted in chapter 6 of this thesis.

Evaluation studies of government assisted training in Australia have mostly utilised aggregate measures such as placement rates. Econometric studies, on the other hand, have tended to focus on the provision of training in the private sector. Reforms in the last decade have increased the scale of government support for the employer-led Traineeships, which now share many of the features of Youth Training in the UK. Considerable effort has been given to assisting target client groups particularly Aborigines and Torres Straits Islanders with a corresponding evaluation effort to match. The latter have confirmed the relative disadvantage facing many Indigenous Australians in the labour market but have suggested that training programmes can reduce this disadvantage.

The potential exists to apply similar evaluation methods in each country but this will require a greater emphasis on training data in the Australian surveys. The next chapter presents the results of a 'Meta-Analysis' conducted on the UK studies surveyed in this chapter. The considerable variation across studies observed in impact estimates between programme participants and non-participants, in terms of both employment probability and earnings differentials, lends itself to meta-analysis. Meta-analysis is designed to provide a more objective identification of the sources of variance in estimates across similar studies. The aim is to achieve greater accuracy than that obtainable from a single study and also to allow the investigation of factors responsible for between-study variation. The chapter includes an overview of the development of meta-analysis and rcvicw of existing meta-analyses within the economics literature.

Chapter 5: A Meta-Analysis of the Impact of Youth Training in the United Kingdom

5.1: Introduction:

The previous chapter illustrated that existing studies investigating the impact of government sponsored youth training in the United Kingdom (UK) have, in spite of applying roughly similar procedures, reported a diverse range of estimated outcomes.[114] This chapter presents the results of a *meta-analysis* conducted on these studies in order to explain the variation among estimates. Unlike traditional literature review methods in which the results of different studies are compared by means of verbal arguments, meta-analysis involves the application of statistical rules to help determine what the literature says in an impartial manner. As such, this approach allows the reviewer to identify any regular patterns discernible in a body of studies reporting diverse outcomes on a given topic and produce estimates of the effects of variation in study method or other study characteristics on the reported empirical results (Jarrell and Stanley, 1990, p. 56). It uses the summary statistics from the original studies being reviewed as the data points for the analysis based on the assumption that each study provides a differing estimate of the underlying relationship within the population. As such, meta-analysis is a means to more accurately summarise research as it is reported and to more efficiently extract useful information from the literature.

Most empirical sciences have developed some form of quantitative research synthesis, or meta-analysis, of replicated research. In particular, meta-analytical methods have been widely applied in the natural and medical sciences but its uptake in the economics profession has been relatively slow. Credit for coining the term meta-analysis is usually attributed to G. V. Glass (1976, 1977) although meta-analytical techniques were already in relatively wide spread use, particularly in agricultural research. Glass, who developed many of the procedures currently

---

[114] The focus on the UK in this chapter is due to the relatively larger pool of available youth training impact estimates available in the evaluation literature.

employed in the social sciences, argued that the same scientific rigour should be applied to literature reviews as to individual empirically designed studies to study the research question at hand (Glass, McGaw and Smith, 1981, p. 12). Glass distinguished meta-analysis from primary and secondary analysis. He defined primary analysis as the original research, which includes data collection, data processing, and publication of results. Secondary analysis was defined as the re-analysis of data following the same research question from either a different perspective or with different techniques. Whereas, meta-analysis draws upon the summary statistics of a multitude of studies in order to integrate the findings without having access to the original data (Glass, 1976, p. 3).

Meta-analysis can be defined as "… essentially the enterprise of combining and reconciling the results of replicated studies" (Hedges, 1997, p. 113) or, more simply, "… the statistical analysis of statistical analyses" (Glass, 1976, p. 3). It provides a framework within which to integrate and empirically explain the various results in a given domain of research and seeks to explain the variation in reported estimates. As observed above, it also allows us to analyse various aspects of the research process itself as, unlike traditional (narrative) review methods, the data points for meta-analysis come from the original studies themselves. Moreover, meta-analysis also possesses the potential to facilitate the transfer of existing knowledge to similar situations, which have not previously been investigated, saving on comparatively expensive research efforts (Baaijens, and Nijkamp, 1997, pp. 1-2).

Wolf (1986, p. 10) lists a number of potential problems with traditional literature reviews that may be addressed through meta-analysis including: (1) the selective inclusion of studies (2) differential selective weighting of studies in the interpretation of a set of findings (3) misleading interpretations of study findings (4) failure to examine characteristics of the studies as potential explanations for disparate or consistent results across studies and (5) failure to examine moderating variables in the relationship under examination. In summary, procedures employed in meta-analysis attempt to reduce the subjectivity involved in the synthesis of the pool of research literature on a particular issue permitting quantitative reviews. Consideration is thus made of the relative merits of methods used across studies rather than the merits of the individual studies themselves. Stanley and Jarrell (1989, p. 134) emphasise this latter point:

> "[meta-analysis] largely avoids the implicit valuation of researchers' work and [by implication] of their reputations that inevitably characterises conventional literature reviews".

The next section presents a brief overview of the development of meta-analytical methods followed by a survey of its application in the field of economics. Section 5.4 then introduces meta-regression analysis as the appropriate basis for the integration of reported findings investigating the impact of youth training. Studies from the UK are then used to illustrate the application and potential of meta-analysis in section 5.5, which includes a description of the data and model. The results of the analysis are then presented and interpreted in section 5.6 while the final section provides a chapter summary.

5.2: The Development of Meta-Analysis:

The origins of meta-analysis can be traced back to the 1930s when agricultural researchers developed procedures for integrating statistics across a set of studies. The innovation of these methods largely stemmed from the desire to combine evidence from a large number of independent agricultural experiments, all of which were planned to test a common hypothesis. Interpretation of the individual experimental findings was hampered by uncontrolled factors arising from differences in soil quality, agricultural practice, climate, and other environmental factors (Wolf, 1986, p. 13). Interpretation was often further compromised by land constraints requiring the experiments to be spread across different sites introducing further uncontrolled factors (Bangert-Drowns, 1987, p. 656).

Two distinct approaches were adopted to resolve the interpretation problem. In the first approach methods were developed by Fisher (1932), K. Pearson (1933) and E. S. Pearson (1938) among others, to test the statistical significance of results combined from the separate experiments. Often labelled the *combined test* approach, methods range from simple summation procedures to more elaborate procedures involving the combining of the raw or weighted test statistics reported in the original studies. The second approach, developed by Cochran (1937) and Yates and Cochran

(1938), estimated the mean effect and variability of a particular treatment examined at different agricultural centres. The second approach is labelled as 'effect size' or *combined estimation*.

5.2.1: Combined Tests

The problems faced by the agricultural researchers in the 1930s when attempting to interpret a large body of related research is not unlike those facing researchers in the social sciences today. The use of combined tests as a viable method for statistically integrating social science research was first advocated by Jones and Fiske (1953). They argued that the methods of combining probabilities developed by Fisher, K and E. S. Pearson could be used to combine findings in education and psychology (Bangert-Drowns, 1987, p. 656).

In general, a combined test is a way of combining statistical significance values (*p* values) from a set of studies that test the same hypothesis, but use different designs or measure outcome variables differently, to obtain an overall level of significance for those studies. Combined tests allow the researcher to determine if the data are consistent with the null hypothesis in all studies or if the data suggests that the null hypothesis is false in at least some of the studies (Hedges, 1994, p. 8).

A more recent version of the combined test, Winer's (1971) method of combining independent *t* statistics, illustrates this approach. In common with most of the alternative combined tests, the aim is to combine individual significance tests from a number of studies into one overall pooled test. For Winer's test, the *t* statistics associated with individual probability values are summed and divided by the square root of the sum of the degrees of freedom (*df*) after each has been divided by itself minus two (*df* - 2), which can be written as:

$$Z_w = \frac{\sum t}{\sqrt{\sum [df/(df-2)]}} \tag{5.1}$$

This procedure is based on $[df/(df-2)]$ being the variance of a *t* distribution, which is approximately normally distributed when $df \geq 10$ (Winer, 1971, p. 50). The procedure yields an estimate (Z_w) of the likelihood that the obtained set of probabilities could have arisen by chance given a true compound null hypothesis

(Strube and Miller, 1986). While relatively straightforward to compute, this procedure is inappropriate when tests are based on small samples (less than 10) or very large samples, in which case $[df/(df-2)]$ would approach unity. To avoid the problems that arise with the latter case, an alternative procedure, the 'Stouffer test' may be used (Wolf, 1986, p. 20). This test first requires each p value to be converted to a standard normal deviate (z_i) with sign maintained to indicate direction of support. These individual z_i s are then summed and divided by the square root of the number of tests combined (N) to produce an overall standard normal deviate (Z_s):

$$Z_s = \frac{\sum z_i}{\sqrt{N}} \tag{5.2}$$

This procedure is based on the sum of normal deviates being itself a normal deviate, with the variance equal to the number of observations summed. There are a wide variety of alternative tests of combined significance available, see for example Rosenthal (1984, chapter 5) and Wolf (1986, chapter 2).

One of the advantages cited for the combined test approach lies in the increased statistical power of the overall comparison.[115] If, for example, several tests consistently favour the research question but fail to reach the level of significance, due to small sample sizes, the overall test would more easily become significant because the pooled sample size is much larger. On the other hand, when a large number of studies are combined, caution has to be exercised as sample size may become artificially inflated and a highly significant overall test would thus not provide useful information.

Procedures that combine significance tests have been criticised for several reasons. First, the focus is on an overall probability instead of on distributions. In other words, if the meta-analysis is to achieve its objective, it is necessary to examine the variation of results because both positive and negative outcomes could cancel each other out, and because conflicting results should represent a challenge to identify

---

[115] For a comparison of the relative statistical power of a range of combined tests, using simulation procedures, see Strube and Miller (1986). They note that most of the tests provide comparable power when applied to a large number of studies but that the procedures vary in their ease of computation making some more preferable under certain conditions.

substantial sources of variation. Various assessments of variation have been proposed to overcome this criticism such as the test of homogeneity suggested by Rosenthal (1984, pp. 76-77).

Second, it is possible that only a few significant results are actually published and many more may remain in 'file drawers' leading to a biased sample of all available studies. This problem also applies to narrative reviews however, while it is not possible to estimate how big an issue it is in meta-analysis, it is possible to estimate the number of additional studies confirming the null hypothesis that would be required to reverse a conclusion drawn from the meta-analysis. This procedure provides some estimate of the robustness and validity of the findings (Rosenthal, 1984; Hunter and Schmidt, 1990).

Third, a reported p value just indicates the probability of an error in rejecting the null hypothesis if it were true (type 1 error). It does not, however, provide an estimate of the magnitude of treatment effects. To resolve this issue we need to turn to combined estimation methods.

5.2.2: Combined Estimation

The combined estimation approach involves measures of *effect size* and it is based on developing standardised, scale equivalent indices of the magnitude of effects, that are independent of the various scales of measurement used in the original studies (Wolf, 1986, p. 13). Before this approach could be adopted in the social sciences, however, two methodological barriers had to be overcome. Specifically, dissimilar dependent measures and variations in study design precluded wider adoption of the meta-analytic method (Bangert-Drowns, 1987, pp. 656). That is, even though a group of studies might investigate the same question and employ similar methods, differences in dependent measures and variations in study specification typically distort any combined estimation results.

The problem was addressed, albeit indirectly, by Cohen in the 1960s. He demonstrated that it was possible to statistically summarise a diverse literature without taking recourse to counting or combining significance levels. The focus of Cohen's work was on statistical power analysis and how to estimate the optimal sample size of a study. He highlighted the potential of effect size in statistical research integration and suggested a rudimentary form of how it might be measured. The latter exposition and application of meta-analysis by Glass and associates built on the work of Cohen (Bangert-Drowns, 1987, p 658).

Glass coined the term meta-analysis and applied a modified form of Cohen's effect size to a collection of studies on psychotherapy effectiveness and another on studies investigating the influence of class size on educational outcomes.[116] Subsequently, most quantitative analyses of outcomes of studies with a similar research question are called meta-analysis.

A key assumption made by Glass is that each study provides a differing estimate of the underlying relationship within the population (Glass, 1977). The key to this approach is the derivation of a standard measure of empirical effect that can be assumed constant across the literature being examined. Effect size as proposed by Glass can be written as:

$$g = \frac{(\mu_e - \mu_c)}{\sigma} \tag{5.3}$$

where μ_e is the experimental group, μ_c is the mean of the control group, and σ is the standard deviation of the control group. Effect size renders the results of highly individualised studies, concerning some common phenomenon, comparable and thus suitable for analysis (Stanley, 1989, p. 163).

Another characteristic of the approach adopted by Glass lies in the inquiry of study features. Instead of simply averaging all available summary statistics, more emphasis was put on relating study features to outcomes and discovering the influence of specific factors, inherent in the research design, on the resulting effect sizes. If, for example, females and males differ with respect to their mean effect size, gender is identified as a moderator. This approach requires a coding scheme where as many study features as possible are entered in order to examine their potential moderating influence on effect sizes (Glass, 1977). He also recommended that,

---

[116] The results of a number of meta-analyses on different approaches to estimating the impact of class size on educational outcomes are reported in Glass and Smith (1979).

rather than integrating all summary statistics on the whole, it is preferred to recognise subsets of data that produce distinctive mean effect sizes and, thus proliferate additional knowledge about the research domain that was not apparent before applying meta-analysis. For example, study features such as age, source of publication, publication year, and technical quality of the study, could be used. Glass et al. (1977, 1981) suggested that these variables should be regressed on effect sizes in order to examine their interrelationships.

Not all empirical data can be readily converted into measures of effect size and subsequently there are now a variety of different procedures available for conducting a meta-analysis employing a range of different common statistics. Of particular significance to economics is 'meta-regression analysis', that is, a multiple regression analysis of the results of previously published and unpublished multiple regression analyses. This approach is discussed further below, other approaches are summarised in Hunter and Schmidt (1990, chapter 11, see also Bangert-Drowns, 1987, pp. 657-663).[117] The key feature shared by all the different approaches, however, is the application of a standard measure or 'common metric' to the studies of their respective samples in order to allow integration of the results.

The development and ongoing research into such measures has also allowed proponents of meta-analysis to partly rebut the most common criticism they face, the 'apples and oranges' argument. This argument claims that logical conclusions cannot be drawn by comparing and aggregating studies that include different measuring techniques, definitions of variables (for example, treatments and outcomes) and subjects because they are too dissimilar. This obstacle can however be ameliorated if the studies idiosyncratic characteristics are appropriately coded and statistically tested to determine whether they are related to the meta-analytic results (Wolf, 1986, p. 14-15; Hunter and Schmidt, 1990, chapter 12).

In summary, the combined estimation method of meta-analysis is designed to develop unit-free measures of treatment effects and hence provide a relatively comprehensive method to estimate the average effect of treatments on outcome

---

[117] For an empirical assessment of different meta-analytic methods see Rosenthal (1984, chapter 3).

variables across a large number of studies. Combined tests are generally seen as inferior to combined estimation methods but can be usefully applied if probabilities are the only measure available to the meta-analyst (Hunter and Schmidt, 1990; Wolf, 1986). Alternatively, the analysts may choose to report the results of both combined tests and combined estimation, as illustrated by Hedges, Laine and Greenwald (1994), discussed below.

5.3: Meta-Analyses in Economics:

There are relatively few examples of the application of meta-analysis in economics, although among these topics in labour economics appear to be fairly well represented. Examples are, Jarrell and Stanleys' (1990) study of the wage gap between unionised and non-unionised workers; Hedges, Laine and Greenwalds' (1994) study of differential school (financial) inputs on student performance outcomes; and Card and Kruegers' (1995, Chapter 6) meta-analysis of the minimum wage literature.[118] A brief review of this work serves to illustrate the diversity of techniques and potential applications for meta-analysis within economics and related disciplines.[119]

In an attempt to explain the large variation of the union/non-union wage gap, over time and among different researchers, Jarrell and Stanley (1990, p. 66) used meta-regression analysis to re-analyse and extend the 1986 study by H. G. Lewis. Their analysis explicitly allows for the impact of the business cycle and controls for individual study characteristics. As a result of these adjustments the annual union/non-union wage gap, which Lewis reported as ranging from 9.6 to 16.4 percent, is generally revised downwards to a range of 8.9 to 12.4 percent.

They found a number of different study characteristics that led to a significant bias in estimated results including non-representative sampling. These included, the over-weighting of black workers, the omission of workers under 25 years of age, the use

---

[118] See also the editorial note in the 1997 (4th quarter) issue of *Labour Economics* (Arulampalam, Hartog, MaCurdy, and Theeuwes, 1997, pp. 99-105) encouraging the submission of replication and reanalysis studies - which include meta-analyses.

[119] For a more detailed review of the application and practice of meta-analysis in economics, focusing on environmental economics, see Bergh (van den), Button, Nijkamp and Pepping (1997).

of different wage measures (hourly, weekly or annual), whether wage measures were converted into natural logarithms or not, and the omission of certain independent variables (including industry and extent of unionism). Jarrell and Stanley also found that authors who published more than one estimate of the union/non-union wage gap nearly all produced estimates below the average estimate reported by Lewis (ibid. pp. 61-64).

In a separate meta-regression analysis Jarrell and Stanley allow for cyclical variation as represented by the unemployment rate. The resulting coefficient is positive and significant indicating that the original studies 'overestimated' the wage gap.[120] In summarising their study, they argue that the application of meta regression analysis results in "...a clearer picture of how the reported wage gap is influenced by a host of study characteristics and economic conditions...", the identification of weaknesses in the literature and a new base line against which future results can be measured (ibid. p. 66).

Baaijens, and Nijkamp (1997) adopt a similar approach to examine a sample of studies on tourist multipliers. These are designed to measure the effects of tourist expenditure on the economy of a region. They apply meta-regression analysis in order to determine on what basis the average multiplier values computed by various researchers across different regions differ from each other. Limited by a small sample they adopt a step-by-step approach, first establishing a base equation then extending this to include other variables in order to test various hypotheses. After establishing a positive relationship between the regional multiplier and population size, they select this as their base and then examine the influence of various features employed in the different studies including the economic model employed and background features (Baaijens, and Nijkamp, pp. 10-11).

Baaijens, and Nijkamp find that multipliers estimated by Archer's model appear to be smaller than those estimated using the input-output model, although a significant

[120] Note that a meta-regression coefficient adjusts from the standard, measured by the intercept, to the individual study's original gap measure. This implies negative coefficients indicate the original study underestimated the gap, and positive coefficients indicate they overestimated the gap (Jarrell and Stanley, 1990, p. 63).

difference is not identified. Amongst the background variables they find that tourist numbers, the tourist to population ratio and the share of the most important country of origin in the total number of tourists all have a positive impact on the tourist multiplier (ibid. pp. 12-13).

They also explore how meta-analysis can be utilised in order to estimate the value of the tourist income multiplier in regions outside those investigated in the original studies. This is achieved by entering the required background characteristics for the country with an unknown multiplier into the regression equations used in the meta-analysis. Using the results of the meta-analysis to choose the appropriate equations, an interval of plausible multiplier values is then derived for the country of interest (op. cit., pp. 14-15). Baaijens, and Nijkamp argue that such transferability of known research findings to a similar situation, which has not previously been investigated, is an area of great potential for meta-analysis (ibid. pp. 1-2). The authors stress however that, because of the small dataset employed, their results are only indicative.

Hedges, Laine and Greenwald (1994) adopted both combined significance tests and combined estimation methods to reassess the findings of E. A. Hanushek's earlier narrative literature review on the relationship between (financial) school inputs and student performance levels. Hanushek (1991) questioned whether there is a relationship between per-pupil expenditure and accomplishment of students in school. Hedges et al. first applied combined tests to this question using a range of resource inputs (facilities, administrative inputs, student/teacher ratios, teacher salary, teacher education, teacher experience, and per pupil expenditure). The results showed a statistically significant relationship between all the inputs and student achievement, with the exception of school facilities (Hedges et al. 1994, pp. 9-11).

These findings were subsequently supported by a meta-regression analysis applied to the per-pupil expenditure and teacher experience variables. Teacher salary was also generally positive but the magnitude was too small to be of practical importance. Teacher education reported a negative outcome for each sample of studies, while the remaining variables showed a mixed pattern of regression coefficients. Hedges et al. made the observation that "...resources matter, but allocation of resources to a specific area (such as reducing class size or improving facilities) may not be helpful in all situations. That is, local circumstances may determine which resource inputs are most effective..." (ibid. p. 11). After reconciling some of the apparent contradictions between the combined test and meta-regression approaches, they conclude that Hanushek's findings (accepting the null hypothesis) are not supported by the particular studies he sought to synthesise in his review (op. cit., p. 13).

Smith and Huang (1995) examine 86 reported estimates from 37 different studies on the marginal willingness to pay for reducing particulate matter from hedonic property value models developed between 1967 and 1988. They identified a consistent relationship between these estimates, the level of air pollution in each city and the average income of its residents as well as a number of other variables (ibid. pp. 216-220). They extended their study to test the effectiveness of meta-analysis summaries in approximating marginal valuation estimates in the face of changing circumstances. Smith and Huang conclude that meta-summaries reduce sensitivity to extreme results when compared to the original hedonic models and may serve a role in estimating the impact of policy changes when existing results are not directly relevant to the questions or geographic locations involved (ibid. pp. 223-224).

Card and Krueger (1995) apply meta-analysis in order to determine whether published time-series studies, reporting adverse employment effects of the minimum wage, have been affected by 'publication bias'. In other words, they sought to directly address the file drawer issue raised above as a criticism in relation to combined tests. Their inquiries began by assuming publication bias induces a tendency towards the reporting of t-statistics that exceed 2, they then examined the reported t-statistics and plot the relationship between these and sample size. Contrary to expectations, based on statistical theory, the plot showed a negative relationship between the value of the reported *t*-statistics and sample size suggesting the possibility of publication bias (op. cit., pp. 187-188). Card and Krueger use a form of meta-regression analysis and find that the negative relationship remains robust in the presence of other variables. Publication bias is only one possible explanation for the outcome reached and they acknowledge it may also be due to specification searching or structural change in the models employed (ibid. p. 205).

Meta-analysis techniques are applied to a collection of studies investigating the aggregate demand for labour in Australia by Doucouliagos (1997). Applying meta-regression analysis, his results revealed a positive association between demand and aggregate employment and a relatively greater inverse relationship between the real wage and aggregate employment. He found that, with the exception of the association between demand and aggregate employment, 80 percent of the between study variation could be attributed to specification differences, estimation procedures, the frequency of the data and sample size. Doucouliagos notes, however, that meta-analysis could not make up for the deficiencies in the original studies that failed to establish causality and which contributed little to help determine what the correct specification should be. Moreover, he could not rule out spurious regression in the bulk of the studies and noted that most failed to test for asymmetry.[121] Consequently, Doucouliagos concluded that the available pool of aggregate labour demand studies was of little use for policy and recommended that a new batch of studies need to be undertaken urgently (1997, p. 237).

The meta-analytical studies surveyed above illustrate the versatility of meta-analysis and its potential application in economics. Most of the analyses reported tests on a range of different specifications and subordinate data sets to assess the robustness of their results. For example, Hedges et al. (1994, p. 8) compare the results of their full-sample analysis to a trimmed sample, with the largest 5 percent and smallest 5 percent of reported estimates removed, in order to limit the influence of deviant effects or 'outliers'. Similarly, they allowed for the disproportionate influence of single articles by including a sample minus the estimate associated with the article which, when excluded, would impact most on the overall results. Meta-analysis can thus be viewed as a complement to traditional review methods by statistically explaining differences between the results of similar studies. Moreover, it is also an effective means of consolidating the experiences gained from existing studies and in determining useful directions for future research.

---

[121] For example, if it is found that increases in the real wage do contribute to unemployment it does not necessarily follow that a cut in the real wage is required to reduce unemployment (Doucouliagos, 1997, p. 237).

5.4: Meta-Regression Analysis

Given any empirical literature review in economics involves the summarising and evaluation of regression results, typically ordinary least squares estimates, Stanley and Jarrell (1989), propose that meta-regression analysis can be usefully employed to integrate the literature. Empirical studies, suitable for meta-regression analysis, are those that attempt to either identify the determinants of economic phenomena, to estimate the magnitude of the interconnections among economic phenomena, or to test a particular hypothesis. That is, those investigations that were originally undertaken for explanatory purposes and the critical issue relates to a question of magnitude or significance of some regression coefficient. After all the available primary research coefficients are collated, they become the dependent variable for the meta-analysis. Stanley and Jarrell (1989) summarise the potential of meta-regression analysis as providing "... us with the means to analyse, estimate and discount, when appropriate, the influence of alternative model specification and specification searches" (p. 162). They later, in Jarrell and Stanley (1990), provide a framework for such analysis in applying the method to reanalyse Lewis' renowned study of the union/non-union wage gap. A similar methodology can also be found in Baaijens and Nijkamp (1997) and Doucouliagos (1997).

The meta-regression model expresses the functional relationship between the estimated regression coefficient $\left(\hat{b}_j\right)$ from each of L primary studies and K 'meta-independent' variables $\left(Z_{jk}\right)$ that describe important data and model specifications of an original study. The meta-regression analysis model can then be written as:

$$\hat{b}_j = \alpha_0 + \sum_{k=1}^{K} \alpha_j Z_{jk} + \varepsilon_j \qquad (j=1, 2, ..., L) \qquad (5.4a)$$

where the α 's are the coefficients for the meta-regression variables and ε_j is the meta-regression disturbance term. The estimated α 's can then be examined and interpreted as average effects introduced by the individual specifications of the original studies on the estimated regression coefficient $\hat{b}_j$ observed in those studies. In other words, the meta-regression model should include a number of relevant variables related to the model specification, dataset, estimation and authorship in an attempt to account for any systematic variation attributable to these factors. It is possible to include all relevant studies as long as similar types of numerical results are reported.

As in the classical regression case, the final specification of the meta-regression equation should be determined by reference to the data, in this case the available empirical literature base. For example, there is no general agreement in the meta-analysis literature about whether multiple estimates from a single study should be counted as different observations (Glass et al. 1981, pp. 197-216). Nor on the extent that multiple results from a single author or group of authors vary from the standard or from most other studies. For example, Jarrell and Stanley (1990, p. 58) include a dummy variable for any author(s) with multiple studies in order to control for any effect peculiar to the author(s). In contrast, Doucouliagos (1997, p. 228) chose to average multiple estimates originating from the same author(s).[122] Similar approaches can be adopted to allow for different databases, programme changes and time periods. In this manner an increased quantity of information can be synthesised highlighting one of the central advantages of applying meta-regression analysis over the conventional approach to literature reviews.

Stanley and Jarrell (1989, pp. 166-167), note that since the original studies use different datasets, with varying sample sizes and different independent variables there is a strong likelihood that the variances of the estimated coefficients may not be equal. Hence the meta-regression errors may be heteroskedastic. They suggest that this may corrected for by dividing the meta-regression equation by the standard error S_b for each coefficient b_j, that is:

$$t_j = \frac{b_j}{S_b} = \frac{\alpha_0}{S_b} + \sum_{k=1}^{K} \frac{\alpha_j Z_{jk}}{S_b} + \frac{\varepsilon_j}{S_b} \qquad (j=1,\ 2,\ \ldots,\ L) \qquad (5.4b)$$

The potential heteroskedasticity directs the emphasis of meta-regression analysis to the reported t statistics, which has intuitive appeal because they generally provide the

[122] Doucouliagos was, however, mainly concerned with cases where statistical independence may be jeopardised if the same author uses the same dataset but experiments with different modelling or estimation procedures, that is 'conceptual replication' (ibid. p. 228).

critical tests in the original studies (Stanley and Jarrell, 1989, p. 166; see also Doucouliagos, 1997, pp. 226-227). Similarly, Baaijens and Nijkamp (1997) also note the potential source of error arising from the fact that the reported effects in the original studies are only estimates but suggest a second error term can be added to the regression to take this into account. In either case, a potential limiting factor to such corrections is however, that most studies do not report all the relevant statistics and, because of this information constraint, it may be more efficient to assume that the studies have the same quality (Baaijens and Nijkamp, 1997, p. 9). That is, a meta-regression in the form of equation (5.4a) is usually estimated.[123] Baaijens and Nijkamp caution that the consequence of ignoring the actual nature of the individual studies will, however, result in the estimated coefficients of the meta-regression having non-minimal variances.

5.5: Data and Model

Research on the impact of youth training in the United Kingdom lends itself to meta-analysis as it has resulted in a set of wide ranging impact estimates, in terms of both employment probability and wage differentials, between programme participants and non-participants. Further, as observed in chapter 4 (section 4.3), researchers have used a range of different datasets, employing increasingly complex models and more sophisticated statistical techniques in the pursuit of greater precision to assess the impact of training for various programmes. At the same time, both the datasets and the programmes share many key features with the latter having been characterised as different evolutions of the same programme. The evaluation studies surveyed in chapter 4 are summarised in table 5.1 in terms of the key reported impact estimates of the training. Estimates on the probability of being in employment, following participation in a youth training programme, range from negative but insignificant effects to plus 21 percent. The estimated impact of programme participation on earnings across the different studies ranges from minus 16 to plus 12 percent. However, when only the statistically significant estimates are considered the earnings

[123] Amongst the individual studies we include in the meta-analysis only one, O'Higgins (1994), reported potential problems with heteroskedasticity. His corrected estimates are used in the meta-regression.

impact is predominately negative. In this section meta-analysis is hence used in an attempt to explain the observed between study variation given that the studies have used substantially similar datasets on seemingly similar programmes.

By accumulating all the reported outcomes across the studies, one can gain a more accurate representation than is provided by the individual study estimates through the cross validation of coefficient estimates (Stanley and Jarrell, 1989, p. 162). In this case, by adopting Bradley's (1995, p. 33) depiction of modern youth training in the UK as being an evolutionary process, from the Youth Opportunities Programme (YOP) through to Youth Training (YT), we can justify including most of the studies outlined in chapter 4.[124] In terms of impact on employment, the relevant studies number 8, yielding up to 22 estimates across a range of different specifications and sub-samples, similar to the number used in the meta-analyses of Doucouliagos (1997) and Card and Krueger (1995).[125] Eight of the studies in table 5.1 also provide earnings outcomes, yielding a maximum of just 16 estimates. However, Baaijens and Nijkamp (1997) who used 9 studies, with only 11 estimates to draw on, faced similar constraints when applying meta-analysis to the literature on tourist multipliers.

[124] The duration analyses discussed in chapter 4 have been excluded because it is not possible to adjust the duration estimates to a common employment metric.

[125] Doucouliagos (1997) had a total pool of 34 studies to draw on but he used only 21 of these (yielding 15-18 estimates) in the meta-analysis of employment, real wages and demand, and used 13 (yielding 10-12 estimates) in the meta-analysis of unemployment, real wages and demand. Card and Krueger (1995) utilised 15 studies yielding 14 estimates after omitting one study considered an outlier.

Table 5.1: Summary of Studies on Youth Training in the UK by Estimated Impact: [1]

No.	*Author*		*Employment*	*Earnings*
1	Main & Raffe (1983)	*male*	0.058 (1.22)	-
		female	0.142 (2.88)	-
2	Main (1985)	*male*	0.044 (2.24)	-
		female	0.081 (3.73)	-
3	Main (1987)	*average*	-	-0.090 (-1.13)
4	Hutchinson & Church (1989)	*aged 16*	-	-0.086 (-2.79)
		aged 17	-	-0.036 (-1.34)
		aged 18	-	0.004 (0.13)
		adult	-	-0.073 (-3.39)
5	Main & Shelly (1990)	*average*	0.079 (0.75)[#]	-0.067 (-1.91)
		advantaged	0.17 (2.70)	-
		disadvantaged	0.11 (2.75)	-
6	Whitfield & Bourlakis (1991)	*average*	0.181 (2.55)	- 0.03
7	Main (1991)	*average*	0.068 (0.72)[#]	-
		male	0.171 (1.35)	-
		female	-0.084 (0.58)	-
8	O'Higgins (1994)	*average*	0.21 (1.97)	-
		advantaged	0.01 (0.26)	-
		disadvantaged	0.09 (1.51)	-
9	Dolton, Makepeace & Treble. (1994a)	*male*	-	0.047**
		female	-	-0.045*
10	Dolton, Makepeace & Treble (1994b)	*male: average*	-	-0.15
		advantaged	-	-0.014
		disadvantaged	-	-0.127
		female: average	-	-0.16
		advantaged	-	-0.164
		disadvantaged	-	-0.039
11	Dolton, Makepeace & Treble. (1994c)[2]	*W & B average*	-0.132 (0.05)	-
		" *male*	-0.347 (0.08)	-
		" *female*	0.008 (0.06)	-
		M & S average	-0.195 (0.05)	-
		" *male*	-0.446 (0.08)	-
		" *female*	-0.034 (0.06)	-
12	Payne (1995)	*male*	-0.04	-0.08 (-3.03)
		female	-0.03	-0.15 (-6.08)
13	Green, Hoskins & Montgomery (1996)	*average*	-	-0.087

Figures in parentheses are the reported t statistics

* *Reported as statistically significant at the 5% level but no t statistic provided*

** *Reported as statistically significant at the 10% level but no t statistic provided*

W & B: approximate Whitfield and Bourlakis specifications

M & S: approximate Main and Shelley specifications

\# *Completed YTS*

Notes: 1. Further information on these studies can be found in tables 4.1, 4.3 and 4.4 in chapter 4, including detail on explanatory variables, database and model employed.

2. *These estimates come from the working paper with the same title but they were not included in the published article, see* Dolton et al., (1992) "The Youth Training Scheme and the School to Work Transition", *Discussion Paper*, No. 92/3, Department of Economics, University of Hull, Hull.

Unfortunately the outcome estimates reported in the econometric studies of youth training in the UK do not provide as good an environment for meta-analysis as the 37 more readily comparable studies of property values (providing 86 estimates) used by Smith and Huang (1995). Jarrell and Stanley (1990) had an even more numerous collection in examining the union wage gap with 114 studies at their disposal (yielding 152 estimates), which also had the advantage of having been adjusted to more comparable measures in earlier work by Lewis (1986). Nevertheless, there are enough studies to demonstrate the potential of meta-analysis in the field of evaluation research through the validation or rejection of the possible reasons for the observed between study variation drawn from the traditional literature review conducted in chapter 4. These included the application of different models, model specification the inclusion of (or controlling for) different human capital variables.

The observed variation between the studies, listed in table 5.1, in terms of model structure, measurement and data sources however, complicate the configuration of data for meta-analysis. For example, studies by Payne (1995) and Green et al., (1995) [study numbers 12 and 13 in table 5.1] both analyse polychotomous choice models and reported estimated earnings for youth training participants in monetary terms. The results from these studies have been converted to the percentage difference in earnings between participants and non-participants to make them comparable to the estimates reported in the study by Dolton et al., (1995) [number 10 in table 5.1].[126] The outcome estimates are reported for a variety of different classifications ranging from mean impact estimates using the whole sample to mean impact estimates for different sub-samples, such as those based on gender. Moreover, the econometric models employed also differ fundamentally in the specification of the explanatory variables, which number from 9 to 22 across the different models with significant variability in content (see tables 4.1 and 4.3 in chapter 4). Given that the outcome is a function of the explanatory variables this highlights one of the key problems facing literature reviewers who want to compare

---

[126] As each of the studies featuring polychotomous choices returns estimates on a range of alternate transition routes to employment, other than participation in youth training, the alternative 'do nothing' was chosen as the route most closely resembling the non-participant route in the bivariate studies.

different studies, that is, "how do we account for different specifications?" Meta-analysis draws on such variability by analysing the effect of specification choices on estimates and thus, "...implicitly involves an assumption that 'true' model specification is not totally idiosyncratic to each situation" (Assmus, Farley, and Lehmann, 1984, p. 66). In other words, meta-analysis explores relative, rather than absolute, differences in specification (Doucouliagos, 1997, p. 229).

Closer examination of the mix of the youth training studies in table 5.1 allows information relevant to the analysis, such as sample size and definition of variables, to be retrieved but it also identifies some of the criticisms levelled at the application of meta-analysis. Perhaps the most damaging criticism that applies when meta-regression analysis is used for evaluation purposes, as noted by Stanley and Jarrell (1989, p. 166) in their analysis of the union/non-union wage gap, arises from the non-experimental nature of the original studies. This may lead to the results becoming dependent on one another, as might the practice of adopting 'trendy' techniques. On the other hand, they argue that this problem poses no greater threat to meta-regression analysis than for the original applications and may even ameliorate the problem as the data is less likely to show time series patterns common in raw economic data (Stanley and Jarrell, 1989, p. 167).

Wolf (1986) notes that meta-analyses may also be criticised because results from 'poorly' designed studies are included with results from 'good' studies. This could be addressed empirically as long as it is possible to code the quality of design, however, Wolf claims that "...there is little evidence to suggest that the magnitude of the effect is related to the worthiness of the design..." (p. 15). Another relevant criticism, given our sample size, is that published research may be biased in favour of significant findings (Card and Krueger (1995); Hunter and Schmidt (1990); and Stanley and Jarrell (1989). Because insignificant findings are rarely published, this in turn can lead to biased meta-analysis results. Wolf suggests several ways to address the potential bias, one is to review results in books, dissertations, and unpublished papers presented at professional meetings and the like, and compare them to the results in published articles. Another approach is to estimate the number of additional studies with non-significant results that would be necessary to reverse a conclusion drawn from the meta-analysis, thus providing some estimate of the robustness and validity of the findings (see discussion on combined tests above).

In summary, the context of research into youth training differs in a variety of ways ranging from the definition of dependent and explanatory variables to geographic location and time frame. While this does cause some limitations in terms of model and data configuration, it also lends itself to the application of meta-analytical methods because the variation of such factors "…is necessary to provide the contextual variability required for meta-analysis" (Assmus et al., 1984, p. 66). Similarly, Doucouliagos (1997) argues that it is this heterogeneity, which necessitates the use of this method and that:

> "The strength of meta-analysis is that it can be used to detect whether differences across studies are real or an outcome of the way in which the study was conducted" (p. 227).

The meta-regression model to be estimated for this study can be written as:

$$\hat{b} = \alpha_0 + \sum_j \beta_j A_j + \sum_m \pi_m S_m + \sum_k \delta_k D_k + \varepsilon \qquad (5.5)$$

where $\hat{b}_j$ is a vector of maximum likelihood estimates (MLE) of the relevant outcome; α, β, π and δ are the meta-regression coefficients; $A_j = 1$ indicates that the study is by author j, $S_m = 1$ indicates that specification or adjustment factor m applies in this study, $D_k = 1$ indicates that data source k is used in the study and ε is the meta-regression disturbance term.

Authors are included if they are responsible for more than one study to control for any effect peculiar to that author (Hedges, 1994; Jarrell and Stanley, 1990). Similarly, the relevant datasets are included to control for differing impacts arising from sampling techniques and regional variation. The remaining explanatory or 'meta-independent' variables are specification factors associated with the different models used in the individual studies. That is, they are included to control for the choices made by the analysts in terms of model selection and specification. Significantly, investigation of the training studies surveyed in chapter 4, revealed considerable between-study variation in specification factors. However, allowing for the relatively small pool of estimates, these factors can be subdivided into groups according to whether they are features of the model employed, characteristics of the sample or general background characteristics. Model features are included to control for the respective research method and may incorporate, for example, the type of model, correction for sample selection and the number of explanatory variables. Similarly, sample characteristics are included to control for any effect peculiar to the sample used for analysis, which may incorporate the size of the sample or the length of period the survey covered. Background variables generally comply with the human capital framework but vary considerably across different studies. These include individual characteristics, such as whether distinction is made on the basis of gender or advantaged/disadvantaged profile, experience of employment (or unemployment) and regional characteristics (such as the local unemployment rate) and/or the version of the programme.

The main aim of this chapter is whether observed differences in the relevant outcome estimates of programme impact can be attributed to the specification factors identified above. The limited pool of estimates in our sample clearly limits the number of independent variables that can be included in the meta-regression and hence it is not possible to control for all the various specification factors.[127] Consequently, the focus is on those factors already highlighted in chapter 4 (sections 4.3 and 4.4) as possibly linked to between study variation in outcomes, particularly the influence of model type and correction for sample selection. Where possible, the general to specific approach has been adopted, that is, the meta-regression models have been estimated using as many independent variables as possible and redundant variables eliminated on the basis of *F*-tests. The specification sub-divisions identified above, model features, sample characteristics and general background characteristics, are also assessed separately to assess their impact on the training measures. However, given the inherent low degrees of freedom, it is also necessary to stress that the resulting estimates consequently only have indicative value and should be interpreted with caution.

---

[127] The small size of the sample also constrains some testing of ordinary regression model assumptions.

A full list of the meta-independent variables considered (dependent on specification) is provided in table 5.2 although some of these are specific to only one of the training programme outcomes. The individual study's sample size, the duration of the survey (in months) and the number of explanatory variables in the individual study are included as quantitative variables but the majority of the meta-independent variables are expressed as dummy variables. The type (version) of training programme in force at the time of the individual study is included to allow for the impact of programme changes as well as introducing a time dependent variable. Consideration is also taken of the fact that some of the reported impact estimates in table 5.1 are specific to distinctions made on the basis of gender and/or advantaged/disadvantaged profiles by including appropriate controls.

Table 5.2: Meta-Independent Variables

A1	Main	*S8*	Gender distinction
A2	Dolton, Makepeace & Treble	*S9*	Advantaged/Disadvantaged distinction
		S10	Exogeneity of experience
S1	Applied OLS	*S11*	Omitted firm size variable
S2	Applied Bivariate Model	*S12*	Omitted unemployment duration
S3	Used Polychotomous Model	*S13*	Youth Opportunities Programme (YOP)
S4	Corrected for sample selection bias	*S14*	1 year Youth Training Scheme (YTS-1)
		S15	2 year Youth Training Scheme (YTS-2)
S5	Sample size	*S16*	Youth Training (YT)
S6	Length of period covered (mths)		
S7	No. of explanatory variables	*D1*	Scottish Young Person's Survey (SYPS)
		D2	Youth Cohort Study (YCS)

Key:
A *indicates authors of more than one study*
S *indicates specification factor*
D *dataset employed*

The list is by no means exhaustive but we are again restricted by our sample size, which severely limits the possible specifications. Given these constraints, it is not possible to estimate how much of the between-study variation in estimates might be due to specific sample selectivity models, only the impact of this class of models in general. Additional factors, external to the evaluation studies, might be considered for inclusion as the pool of available estimates grows, including the state of the economy, as represented by the aggregate unemployment rate, and/or legislative change, such as the introduction of the minimum wage.

5.6: Results

Separate meta-analyses are conducted on each of the evaluation outcomes, the probability of being in post-training employment and earnings potential, as summarised in table 5.1. Clearly, model specifications vary significantly and not all the individual outcome measures are directly comparable, as many are specific to distinctions made on the basis of gender or advantaged/disadvantaged profiles and not all of them are statistically significant. These factors are addressed by estimating the meta-regression using several different sample constructions for each of the respective outcomes. First, all the available observations are included with a selection of relevant variables. Second, only the average impact estimates (over the entire sample) are included and finally, only estimates reported as statistically significant are considered.

5.6.1: Variation in Employment Estimates

The key results from the different meta-regression models, applied to available estimates of training impact on subsequent employment probabilities, in terms of sources of variation, are reported in table 5.3 with the relevant *t*-statistics reported in parentheses. Including all the available estimates allows the most complete specification including most of the relevant independent variables listed in table 5.2. Main (1985, 1991 and also co-author 1983, 1990) is included as an author of multiple studies. The key model features controlled for were whether or not the study made allowance for more than two routes in the transition from school-to-work (polychotomous choice models), whether it corrected for sample selection and the relevant number of explanatory variables. Similarly, sample characteristics include the sample size and the period covered by the study. While the key background variables were whether or not a distinction is made on the basis of gender or advantaged versus disadvantaged circumstances. Other background variables initially considered included the use of labour market experience and firm size variables, the version of the scheme in force at the time of the study and which dataset was employed. The initial sample is comprised of the 22 estimates reported in the eight studies investigating employment that were summarised in table 5.1. These ranged in magnitude from minus 44.6 percent to plus 21 percent with an average estimated impact of training on subsequent employment probability of just over half of one percent.

Table 5.3: Meta-Regression Results: Employment

Variables	*All estimates*	*Averaged Estimates*	*Significant Estimates*
Sample selection model	-0.403 (2.79)*	-0.401 (4.0)	
No. of explanatory variables	0.023 (2.80)	0.029 (3.52)	0.024 (0.43)
Used experience variable	-0.223 (3.43)	-0.221 (3.55)	0.051 (1.04)
YTS-1	0.312 (2.05)	0.303 (3.03)	0.039 (0.76)
constant	0.140 (1.38)	-0.187 (1.95)	0.067 (1.18)
R^2	0.54	0.85	0.51
F-Ratio	4.61	5.49	5.21
N	22	9	7

* *t-statistics are shown in parentheses*

In column one of table 5.3 it can be seen that the results of the meta-regression, when using the full sample, suggest that much of the between study variation in employment estimates is due to the use of sample selectivity models. The results show that studies employing sample selection techniques tended to result in training impact estimates that were significantly lower than the remaining studies, in the order of minus 40 percent relative to the sample mean.[128] Similarly, a sizeable proportion of the variation in reported impact estimates was found to be due to studies that incorporated work experience variables in their analysis, again resulting in estimates that were lower than the sample mean. On the other hand, studies focussing on the one-year version of the youth training scheme (YTS-1) tended to be associated with higher estimates for training impact while additional explanatory variables were also associated with higher estimates, albeit marginal. Surprisingly those studies, which made distinctions on the basis of gender or

[128] The meta-regression coefficient only explains cross-study variation relative to the mean impact of those same studies. It does not, however, imply that the sample selection models are the correct (or incorrect) specification.

advantaged/disadvantaged profiles, contributed little to the between-study variation in reported impact estimates. Variables representing the dataset employed, the use of a firm size variable and authors of multiple studies were similarly found to explain little of the between-study variation and these variables were subsequently omitted from the final stage of the meta-regression on the basis of *F*-tests.

When the meta-regression was conducted using only the nine estimates from studies, which employed all of the available sample, that is those studies that reported average impact, the negative coefficients associated with sample selection techniques and allowing for experience are retained. These results are reported in the second column of table 5.3, which illustrates that the positive coefficients associated with the one-year YTS and additional explanatory variables are likewise retained. The meta-regression in this specification estimates that 85 percent of the between-study variation in reported impact estimates is explained by these four factors, compared with only 54 percent when all 22 estimates were used in the regression. The most likely explanation for this is that the range of average reported impact estimates is narrowed to between minus 20 and plus 21 percent (compared with minus 44.6 and plus 21 percent), with an average impact at just below four and a half percent (compared with 0.5 percent). When the meta-regression is further restricted to including only those estimates that were reported as statistically significant in the original studies the available sample is reduced to only seven observations greatly restricting feasible specifications. The regression was initially limited to the four variables that were not omitted in the earlier meta-regressions on the basis of *F*-tests. The dummy variable indicating whether or not the analyst employed techniques to correct for sample selection was, however, subsequently omitted on the basis of an *F*-test. The results for the remaining variables are nevertheless reported in the third column of table 5.3 where it can be seen that the coefficients are much smaller in magnitude and no longer statistically significant.

In terms of overall outcome measures, the predicted impact of training on subsequent employment probability was positive albeit imprecise. In the case of the meta-regression conducted on all the cross-study estimates, controlling for the various study characteristics described above, the predicted impact was just 0.69 percent

(12.48). While in the meta-regression conducted using only the averaged estimates, the predicted impact was 4.46 percent (12.69).[129]

The results of the meta-regression on the estimated impact of training, in terms of subsequent employment probability, need to be interpreted with some caution due to the limited sample size and hence, at best, are only indicative. Nevertheless, on balance, it would appear that a substantial proportion of the between-model variation in estimates of youth training on subsequent employment probability could be explained by the characteristics of the model employed. Specifically, studies that employ sample selection models and/or incorporate labour market experience variables tend to result in impact estimates substantially different to the cross-study mean. The number of explanatory variables employed in the model provide another factor, albeit less pronounced, contributing to the between-study variation in estimates highlighting the need to exercise due care when specifying the evaluation model. Moreover, as might reasonably be expected, variation can also be attributed to the particular programme in force at the time the primary research is carried out.

5.6.2: Variation in Earnings Estimates

The results from the different meta-regression models, applied to the available estimates of the earnings impact of training, in terms of sources of variation, are reported in table 5.4, with the relevant *t*-statistics reported in parentheses. Most of the variables employed in the meta-analysis of the employment estimates are again utilised with the addition of Dolton, Makepeace and Treble (1994a, 1994b) who, along with Main (1987 and co-author 1991), published multiple studies. The initial sample is comprised of the 18 impact estimates reported in the eight studies investigating earnings that were summarised in table 5.1. The between-study variation is not as severe as in the employment case with a range of minus 16.4 to plus 12.7 percent. The majority of estimates are, however, negative and the resultant mean impact of the sample is approximately minus 6 percent.

---

[129] Figures in parentheses are the respective standard deviations.

As can be seen in column one of table 5.4, the use of sample selectivity models again explains a significant proportion of the between-study variation in estimates. The results of the meta-regression, when using the full sample, indicate that studies which attempt to correct for sample selection again tend to result in lower estimates for the impact of youth training on post-training earnings, in the order of minus 13 percent relative to the cross-study mean. The remaining model features are all omitted from the final regression on the basis of *F*-tests, however in this case, estimates made on the basis of gender are associated with a significant proportion of the between-study variation in estimates. With respect to sample features, studies focussing on the one-year version of the youth training scheme (YTS-1) are again found to be linked to relatively higher estimates of the training impact, while the use of larger samples tends to be associated with marginally lower estimates. The five variables retained in the regression are estimated as explaining more than 60 percent of the between study variation in reported earnings impact estimates.

Table 5.4: Meta-Regression Results: Earnings

Variables	All estimates	Averaged Estimates
Sample selection model	-0.131 (2.33)*	-0.135 (3.72)
Gender distinction	0.141 (3.04)	
YTS-1	0.154 (2.85)	-0.118 (2.74)
Sample size	-0.06E-03 (3.33)	-0.06E-05 (2.22)
constant	-0.015 (0.61)	0.004 (0.10)
R^2	0.62	0.88
F-Ratio	3.24	4.78
N	18	7

* *t-statistics are shown in parentheses*

When the meta-regression was conducted using only the reported estimates from studies, which employed their whole sample, or average impact, the coefficients associated with sample selection correction techniques; the one-year YTS and sample size are very similar to those generated in the full sample meta-regression. Naturally, the gender variable is omitted, given the focus on averaged estimates, but with only seven estimates in the sample the regression has little explanatory power. The results are nevertheless reported in the second column of table 5.4. Only five of the estimates, from the studies summarised in table 5.1, were reported as being statistically significant and hence a meta-regression was not run on this sub-sample.

As in the employment case, the results of the meta-regression on the estimated impact of training, in terms of subsequent earnings, suggest that sample selection models significantly contribute to the between-study variation in estimates reported in the youth training evaluation literature. The analysis again raises the issue of specification, in this case highlighting distinctions made on the basis of gender. Similarly, sample characteristics come to the fore in the earnings case as a possible source of variation in reported training impact estimates. In terms of overall outcome measures, the overall predicted impact of training on subsequent earnings was clearly negative. In the case of the meta-regression including all the cross-study estimates, the predicted impact was minus 7.22 (4.80) percent and minus 9.73 (4.65) percent when only the average estimates were included.[130] Again the results need to be interpreted with appropriate caution due to the limited sample size.

5.7: Chapter Summary

Meta-Analysis involves the quantitative review of reported literature findings and uses the summary statistics from the original studies as the data-points for analysis. The aim is to achieve greater accuracy than that obtainable from a single study and also to allow the statistical identification and investigation of any factors responsible for between-study variation. In this chapter a meta-analysis applied to a set of studies examining the impact of youth training in the UK found that much of the observed between-study variation in estimates might be due to the model employed. For example, sample selection models were found to result in estimates substantially different from the cross-study mean when used to estimate the impact in terms of either employment probability or relative post-training earnings. Because of the limited sample size, the results are only indicative but they clearly illustrate the potential of meta-analysis to enhance literature reviews of empirical studies. Specifically, this is achieved through the development and application of a formal procedure for detecting statistically significant differences across a group of related

---

[130] Figures in parentheses are the respective standard deviations.

studies. As identified in this application these differences may be due to the specific dataset, the author or a variety of specification factors.

This chapter included an overview of the development of meta-analysis in the social sciences and a review of existing meta-analyses in economics to further illustrate the potential and versatility of the meta-analytic approach. Applications range from the reconciliation of diverse estimates to the consolidation of survey findings and may even be used to analyse various aspects of the research process itself. The potential to facilitate the transfer of existing knowledge to similar situations, which have not previously been investigated, saving on comparatively expensive research efforts has also been noted and warrants further research. This chapter has also included discussion on how practitioners have addressed potential problems in meta-analysis in the case of economic surveys. Wider application of meta-analytical techniques will undoubtedly lead to further development of the methodology employed.

Unfortunately, largely because of size, the available pool of evaluation studies in the UK is too limited to allow meta-analysis to establish the influence of specific model types on impact estimates. This issue is investigated further in chapter 6, which continues to examine the causes of between-study variation arising from methodological choices made by the analyst by applying a range of different non-experimental methodologies to a common database. This approach, defined as 'sensitivity analysis', thus controls for any between-study variation that may arise from the analysis of different programmes and/or datasets. It is then possible to investigate the influence of the various assumptions, inherent in the application of different evaluation models, in terms of their impact on estimates of the returns to training. The focus of the sensitivity analysis in chapter 6 is on the post-training earnings of young people. However, several different methods for estimating the returns to training in terms of the impact on the subsequent employment probability of participants are also examined. Drawing on data from the Scottish Young Person's Survey (SYPS) chapter 6 thus presents a detailed 'primary' analysis of the returns to government assisted youth training in Scotland in the early 1990s.

Chapter 6: The Impact of Youth Training in Scotland

6.1: Introduction

The review of the relevant empirical literature in chapter 4 highlighted the possibility that methodological choices made by the analyst may have contributed to the between-study variation observed in impact estimates of government assisted youth training programmes in the United Kingdom. This was supported by the results of a meta-analysis that was conducted on these studies and reported in chapter 5. In the development and application of various evaluation techniques for assessment of youth training some empirical studies, for example Main and Shelly (1990); O'Higgins (1994); and Dolton, Makepeace and Treble (1994b), among others, have raised different economic and statistical factors that might contribute to varying or misleading impact estimates. However, these attempts have typically only addressed a sub-set of possible economic and statistical misspecifications and the insufficient overlap across the different studies does not allow the drawing of any firm conclusions relating to the relevance of these factors. Moreover these studies, and the others reviewed in chapters 4 and 5, utilised different datasets, covered several different evolutions of the youth training programme, and were applied to different regions of the United Kingdom, all of which may have contributed to the observed variation. This chapter extends the exploration of the causes of observed between-study variation by systematically investigating the sensitivity of impact estimates to the application of different non-experimental methods. Furthermore, to control for any variation that may arise from the analysis of different programmes, regions and/or datasets a common database, the Scottish Young Person's Survey, is employed.

The primary focus of the sensitivity analysis is on identifying any variation in estimates of post-programme earnings impact that may be induced by model choice. The review of existing evaluation research in the UK on the earnings impact of youth training, reported in section 4.3.2, revealed that estimates varied between minus 16 percent and plus 4 percent, dependent on the type of individual (and model employed). This variation, in part, indicates the likely presence of sample selection bias in the earnings data. As discussed in chapter 3, selection bias can arise from the fact that we do not observe the decision to participate in training only the realisation of that decision. A variety of models have subsequently been developed to address this problem by taking into account differences between participants and non-participants. The application of different models to the evaluation of youth training in the UK may explain some of the variation in estimates. The key advantage in applying sensitivity analysis to this issue is that it allows the effect of model choice to be evaluated while controlling for various exogenous factors. It should however be stressed that estimating the impact of a programme in terms of earnings is itself subject to a number of limitations. In particular, youth datasets typically only provide a snapshot early in the individual participant's life-cycle earnings. In this analysis the respondents to the Scottish survey have, on average, only attained 19 years of age and many are still in training or education. At the same time, the survey is representative of many available social science datasets and the evaluation of earnings impact does help identify what type of individual gains from programme participation.

This chapter also explores the effect of participation in government assisted youth training in terms of its estimated impact on a programme participant's subsequent employment prospects. The review of evaluation studies in the UK on the employment impact of training, reported in section 4.3.3, revealed that estimates ranged from slightly negative to plus 21 percent. Again part of this variation is likely to be due to model selection and this possibility is investigated through the application of several different models to assess the impact of youth training using SYPS data. Most evaluations of employment impact adopt a probabilistic framework, with the most common evaluation method being the latent variable Probit. The outcome of interest is hence the probability of being in employment at some date after participation in the programme. Sample selection problems might be considered less pervasive in the employment case as there are fewer restrictions on data given that employment status is observed in all cases excepting where respondents have left the labour market. Most studies do, however, exclude those individuals who enter further education on the grounds that they are an inappropriate control group. The typical focus of the existing employment studies has been on estimating the mean impact of training participation on the probability of subsequently gaining employment, which only provides a limited measure of programme impact. Here the analysis is extended through the estimation of alternative impact measures in order to identify the impact for particular participant groups.

The next section of this chapter examines the key features of the Scottish Young Person's Survey exploring its strengths and weaknesses for the purpose of evaluating training programmes. It identifies the respective proportions of respondents choosing different school-to-work transition routes, details the construction of relevant samples and examines the breadth of survey information. Section 6.2 also details how participation in youth training is identified from the survey responses and defines the training variables used in the subsequent analysis.

The remainder of the chapter then proceeds as follows. Section 6.3 reports the estimated impact on earnings of the Scottish cohort following participation in youth training. In section 6.3.1 the methodology employed to compute the dependent variable, ordinary hourly earnings, from the SYPS data is detailed. Following this, the choice of a valid baseline specification is discussed focussing on the different variables that might influence the outcome. In section 6.3.2 the base line specification is established using the orthodox estimation technique of Ordinary Least Squares (OLS). Although the resulting estimates are likely to be subject to sample selection bias, this 'first generation' approach serves as an informative starting point for our analysis. Section 6.3.3 then examines the participation decision, giving particular emphasis to the variables that might help determine whether or not the young person participates in training.

Section 6.3.4 reports the results of the sensitivity analysis applied to assess the effect of employing different non-experimental estimators. The resultant earnings estimates are presented for the different models using the same order as used in section 3.5 (chapter three) where the models were formally outlined. To help communicate the results the relevant estimates are applied to different individual stereotypes in order to determine the impact of participation in terms of the ordinary hourly wage rate. The aim is to reveal more about the construction and interpretation of the resulting estimates and hence aid selection of the appropriate model for the dataset. It is not, however, possible to effectively apply the full range of selectivity models presented in chapter 3 to the Scottish data due to various limitations as outlined in sections 6.3 and 6.4.

Section 6.4 explores the impact of youth training in Scotland in terms post-training employment probability. The variables making up the employment equation are defined in section 6.4.1, which also details how they were identified from the SYPS data and presents the relevant descriptive statistics. The question of variation in estimates is again addressed through employing different estimation models but the focus is on extensions to the standard application of the Probit model. This is followed by a chapter summary in section 6.5.

6.2: Data Description: The Scottish Young Person's Survey

The sample employed in this paper is drawn from the fourth cohort of the Scottish Young Person's Survey (SYPS). The cohort arm of the SYPS surveyed sample year groups, from all secondary schools in Scotland, in the Spring following the end of their compulsory schooling (at the end of their fourth year at secondary school or 'S4') at 16 years plus. This was followed up with another sweep of the survey thirty months later, when the respondents were approximately 19 years old. The Centre for Educational Sociology at the University of Edinburgh administered the survey in conjunction with the Scottish Office Education Department and holds the data in the Scottish Education Data Archive.[131] The cohort arm of the survey began in 1985 with the aim of achieving a nationally representative sample of youth drawn from all Scottish secondary schools. The sample target was 10 percent of the relevant population and the survey instrument was a postal questionnaire. In the follow-up survey the respondents completed a retrospective 'diary' detailing what they had been doing at six monthly intervals since the completion of S4.

The SYPS has covered the experiences of four cohorts between 1985 and 1993, that is, those contacted at 16 plus in Spring 1985, 1987, 1989 and 1991 accompanied by

---

[131] This section draws on the technical reports produced by the *Centre* for each of the respective cohorts.

follow up survey sweeps in Autumn 1987, 1989, 1991 and 1993 respectively. The process of sample construction began the year prior to initial contact of each sweep by contacting all secondary schools in Scotland to request birth dates and other basic details on their pupils. Most of the forms sent to the schools were pre-printed with the names and dates of birth of any pupils that appeared on previous statistical studies thus requiring the schools to simply update the forms as appropriate. The sampling is quasi-random, with samples selected by date of birth to achieve the 10 per cent sample target. For example, students born on the 5th, 15th or 25th might be included in order to achieve the target.

Administered at the same time as the School Leaver's Survey,[132] which makes up the other arm of the SYPS, two questionnaire types were administered in the first sweep in order to make the questionnaires as relevant as possible to respondents. One was supplied to those that stayed on into their fifth year of secondary school (S5) and the other to those that left during or at the end of S4. In addition, and crosscutting questionnaire types, there were two versions of the questionnaire, randomly administered that contained some variation in questions.[133] Some questions covered overlapping ground to enable checks for consistency of responses.

There was some variation in response rates in the different survey sweeps and the issued sample differed from the ideal sample target. Some information received

---

[132] The survey of School Leavers predates the SYPS with records starting in 1962 and collected biennially since 1977. Following a review by the Scottish Office Education and Industry Department in 1991, the survey was re-launched in 1992 in its current format as the 'Scottish School Leaver's Survey' administered by Social and Community Planning Research with the data deposited at the ESRC Data Archive at the University of Essex. Similar to the SYPS, the re-designed Leavers' survey focuses on age/stage cohorts, rather than cohorts of leavers, with a fresh sample of 16-17 year olds approached every two years. The new Leavers' survey also incorporates two follow-up sweeps, the first two years later and the second four years after that (when sample members reach 22-23 years of age). For more detail on this version of the survey see Lynn, P., (1997) "The Scottish School Leavers Survey – an important data source', *Labour Market Trends* (pp. 25-57). Data from the first completed survey was due to be made available for analysis in 2001.

[133] For example, wages received (for those working) and wages anticipated (those not working) were asked for in gross terms, before any deductions such as tax and national insurance, in one version of the questionnaire and the same variables in net terms in the other version. The consistency and validity of responses to this variation of questions, with respect to earnings in the 1985 School Leaver's Survey, has been investigated by Main (1988). He found that respondents that were in employment were able to provide consistent responses concerning gross and net earnings while those who were not employed tended to underestimate net earnings through an inadequate allowance for deductions due to tax and National Insurance.

from the schools was discovered to be inaccurate during or before the survey process, such as returns on students not in the target population or duplicate names. In this case the sample target was reduced to reflect these errors, for example in the fourth cohort the target was reduced from 6,428 to 6,421. Sample members that had moved overseas were not part of the issued sample but were retained as unachieved members of the sample target. In the case of those sample members who appeared to no longer be at the given address, assistance was sought via the careers offices connected with the relevant schools that supplied updated addresses where available. Reminder postcards were also sent to respondents at predetermined intervals as a further method of attrition reduction. The achieved response rates for the first sweep of the 1991 cohort are given in the upper portion of table 1.

The follow-up sweep typically contained three questionnaire types, targeted according to the respondents' status at the time of the first sweep, that is, if they were in full-time education, had left school or whether they did not respond at the time of the first sweep. Only one version of each follow-up questionnaire was used. Attrition reduction methods were similar to those administered prior to the first sweep with forwarding addresses sought from changed occupants and inquiries made through the telephone directory service. Where these attempts to identify sample members were successful, telephone reminders were then utilised in order to help maximise survey returns.

The response to the 1993 follow-up sweep is detailed in the lower portion of table 1, showing that the achieved response rate, minus non-respondents to the first sweep, and allowing for those respondents who were deceased or ineligible, was on average 66.76 per cent. This resulted in an effective overall response rate, for those completing both sweeps of the survey, of approximately 46 percent of the original sample target.

Table 6.1: Response by questionnaire type by number and percentage of sample target.		
Questionnaire type	*number*	*percentage*
School leavers	910	50.61
Staying on	3,540	76.57
Total	**4,450**	**69.30**
Follow-up survey	2,757	66.76
Overall response rate	**2,757**	**46.26**

Source: Centre for Educational Sociology, SYPS Technical Reports

The *Centre* compiled detailed records on reasons for non-response and observed response rates varied across the survey sweeps by gender, educational achievement and time of leaving school. The Centre therefore includes within the dataset various weights, which employ the above variables as weighting factors, in order to better represent the sample population. The weighting method generally involved the construction of a representative matrix for the population and a comparable matrix for the sample. The population figure in each cell was then divided by the sample figure for that cell to yield a grossing factor. These were then multiplied by a constant (the overall 'achieved to population' ratio) to give a final weight. More detail on the specific weightings employed is included in the technical report accompanying each cohort and can be separated from the raw datasets if other sample correction methods are preferred.

The central feature of the SYPS is the diary, which requested respondents to indicate their labour market or educational status for each six-month interval of the survey period. The primary aim of constructing the survey in this way was to collect information on all education, training, jobs and periods of job search the individual had in the relevant period in order to establish a dynamic profile of the individual respondent's experience in and out of the labour market. However, this means that the data are based on the young persons' own reports of their full-time status and as such may generate recall problems. Further, the six monthly observations only capture the 'main' activity as reported by the young person. As the respondent may have changed jobs, moved in and out of training programmes, left school or made some other transition within this time period, the sample may not accurately capture prior labour market status. The wide range of questions across the different survey sweeps does however allow a reasonable degree of crosschecking to be carried out so as to minimise potential problems arising from recall. Furthermore, the structures of the SYPS is fairly typical of most longitudinal or panel datasets available in the social sciences, which provide a succession of snapshots of the respondents' status and activity over a set time period.[134]

The survey also gathered information on a wide range of socio-demographic characteristics from respondents including levels of educational attainment, gender and family background. For example, questions relating to educational attainment asked respondents for the number of O/S grades at level A-C/1-3 attained in S4, as well as details on subsequent study or training courses and any qualifications consequently gained. The survey questions sought to establish family background by, for example, asking the respondents about their parents or guardians' employment status, occupation, qualifications and type of housing.

Thus, in addition, to the practically continuous diary record, there is detailed information about the individual leading up to and including the time of initial contact, as well as at the time of the follow-up surveys. A sizeable portion of the responses are qualitative in nature such as expectations, likes and dislikes which could prove valuable in determining the degree of satisfaction or applicability in an empirical study. Table A3, in the appendix to this thesis, provides a descriptive summary of the main variables collected in the SYPS clearly illustrating the breadth of the survey.

The key economic outcome variables, earnings and employment, can be identified from the responses together with a good number of characteristics common to both participants and non-participants of a particular programme, which allows a variety of econometric models to be constructed and tested in order to assess various

---

[134] The second of the datasets explored in this thesis, the Australian Youth Survey, is based on personal interviews, which are generally considered to be more accurate as it is possible, for example, to encourage the respondent to make reference to a recent payslip.

hypotheses.[135] For example, the effect of gender or region on the likelihood of gaining employment after training or staying in employment once it has been obtained.

The detail in the survey also allows distinct group and/or individual profiles to be compiled and sub-samples fitting these profiles to be assessed relative to the population sample. For example, utilising the School Leavers arm of the SYPS, Main and Shelly (1990, p. 501) distinguish between advantaged and disadvantaged school leavers in their assessment of YTS effectiveness. They define a disadvantaged school leaver as: male, aged less than 18 years, unqualified, with an unemployed father and in an area with an unemployment rate exceeding 21.1 per cent. In contrast, they define an advantaged school leaver as male, aged 18 with four or more O-levels, in full-time work with training in a service industry, having a white-collar occupation and in an area with an unemployment rate of 10.3 percent or less. SYPS data has also been used in the evaluations conducted by Main and Raffe (1983), Main (1987 and 1989), which were discussed in chapter 4 of this thesis.

Another of the aims of the survey was to detail the choice of route in the school-to-work transition from five different options. Other than youth training the choices were further education (those choosing to remain at school); higher education;[136] employment and unemployment. Naturally these routes are not mutually exclusive, Patterson and Raffe (1995) drawing on data from across all four cohorts examined the changing trends in route choice and combination. More recently cross-cohort data from the SYPS has been utilised in the UK submission to the OECD's *Thematic Review of the Transition from Initial Education to Working Life* and for the *Home Internationals* project.[137]

---

[135] The construction of the outcome variables is detailed in the relevant sections, 6.3 and 6.4, below.

[136] No distinction was made between further and higher education in the first cohort (1985-87).

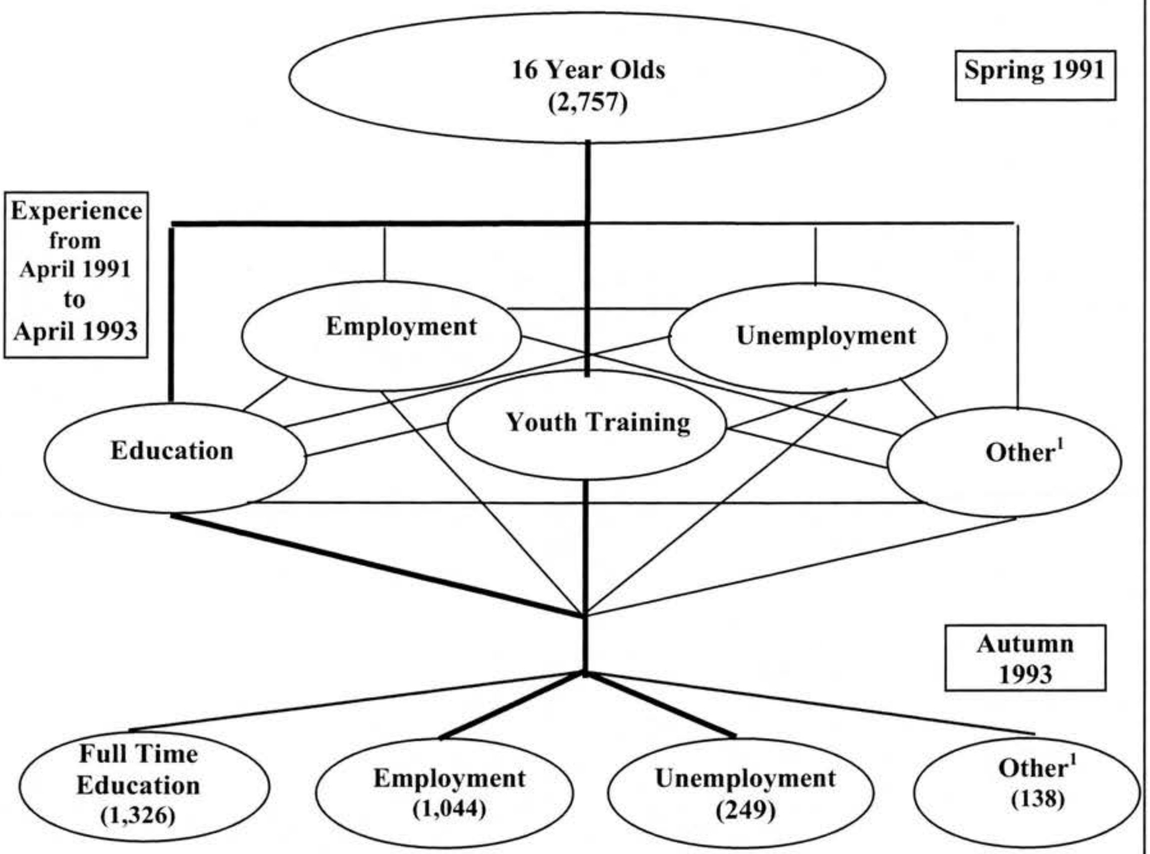
[137] The *Review* was conducted as a follow-up to the OECD's *Jobs Study* conducted in the early 1990s. More detail on the *Review* can be found in the background report by Raffe, D., Biggart, A., Fairgrave, J., Howieson, C., Rodger, J., and Burniston, S (1998) or on the OECD website, http://www.oecd.org, which provides access to a number of country and background reports submitted by participating members. The *Home International Project* was funded by the UK Economic and Social Research Council and reviewed existing statistics and research, interviewed policy-makers in the four territories, and constructed and analysed an integrated dataset on the

Figure 6.1 is a schematic representation of the possible school-to-work transition or further education routes open to school leavers recorded in the fourth cohort of the SYPS with the relevant sample numbers recorded as following each path summarised in parentheses.[138] The evident diversity of transition experience impacts on evaluation as it increases the possible approaches that can be taken in analysing the data. For example, the analyst might choose to focus on specific transition routes such as the two possibilities represented by the bold lines in figure 6.1. The first represents those who opt to leave school at the earliest possible opportunity and progress into the labour market via government supported youth training and second, those who remain at school for a further period before entering the labour market. The second group might be constrained to exclude those who remain at school for more than one year in order to differentiate them from respondents following an academic route. The impact of choosing to enter youth training relative to staying on at school could then be evaluated using a dichotomous choice model. Clearly, a range of alternative approaches of varying complexity can be constructed including the application of polychotomous choice models when comparing more than two routes in the school-to-work transition, see for example Dolton et al., (1994a).

experience and transitions of young people in the 1990s. More detail can be found in Raffe, D., Brannen, K., Croxford, L., and Martin, C. (1999) "Comparing England, Scotland, Wales and Northern Ireland: the case for 'home internationals' in comparative research", *Comparative Education*, 35(1), pp. 9-25.

[138] Table A4, in the appendix to this thesis gives a statistical breakdown of the sample structure.

Figure 6.1: Paths taken by young people reaching minimum school leaving age*



** Figures in parentheses represent sample sizes for those completing both sweeps of the survey*
[1] Other: includes those looking after family, disabled or unable to work.
See also table A4 in the appendix, which details the sample construction.

The starting point adopted in this chapter is a relatively simple comparison between an average participant in youth training with an average non-participant, excluding those who either remain in full-time education (the academic route) or who are outside the labour market. The analysis considers the respondent's position at the time of the follow-up sweep using information from the earlier sweeps largely as explanatory variables. Clearly, the more relevant characteristics that we can control for, in our choice of explanatory variables, the greater the explanatory power of model is likely to be although this cannot prevent the problem of sample selection operating through the training variable(s). Various specification factors are examined in the following analysis reported below but given our focus on the effect of methodological choice on estimates and the characteristics of the SYPS data relatively simple model constructs have been aimed for wherever possible.

As expected a large portion of the survey respondents (1,326) were enrolled in full-time education at 19 years of age. While a further 138 respondents were defined as being outside the labour force because they were looking after family or unable to work. The samples adopted for the analyses reported in this chapter are hence drawn from the remaining categories (employment and unemployment) but vary depending on the outcome of interest being evaluated, earnings or employment. For example, we can only observe earnings for those sample respondents who are actually in employment at the time of the follow-up surveys, which reduces the potential sub-sample to 1,044 cases. The earnings sample is also constrained by missing information on some of the relevant survey questions preventing the computation of 'valid' earnings data in some instances reducing the number of cases to 908 (the computation of the earnings variable is detailed below). It is important to note that that the employed group is likely to differ in several important characteristics relative to the cohort members who are unemployed or out of the labour force. It is also likely that the data constraints on the earnings sample exacerbate both observed and unobserved differences between participants and non-participants in youth training. For these reasons there is a high likelihood that the SYPS earnings data are subject to sample selection bias necessitating the choice of an appropriate selectivity correction model for effective analysis. The effect of this choice is examined in section 6.3 below.

On the other hand, in the employment case, the outcome can be represented as a relatively simple dichotomous variable allowing a broader range of survey data to be utilised in the analysis. The potential sample, in this case, is hence 1293 cases limited only by the responses to particular questions in the survey. The analyst may however impose restrictions on the sample, in order to refine the focus of the analysis and to minimise any pre-existing unobservable differences between participants and non-participants, thus mitigating potential selection problems. For example, the evaluation may concentrate on a certain age group, or those who left school at the end of a certain level of schooling, such as S4 leavers (at approximately 16 years of age). This would hence exclude those individuals following an academic route, reducing the sample to 596 cases, and is generally justified on the grounds that those who choose to stay-on at school are an inappropriate comparison group. Furthermore, using the diary entries, employment status can be determined for various intermediate dates allowing a wider range of controls to be utilised in the analysis. Most studies consequently utilise a probabilistic framework to directly estimate the employment impact of participation in a youth training programme and this is the starting point for our analysis in section 6.4. Variations are possible however, some explicitly controlling for sample selection, and several of these are assessed in section 6.4.

Given several different sample constructs are employed in the evaluation, reported below, the descriptive statistics on the different variables employed are detailed in the relevant sections, 6.3 and 6.4. The remainder of this section discusses the identification of different aspects of youth training participation within the SYPS data and the construction of relevant variables of interest, taking into account the context of government assisted youth training and provision in the UK in the early 1990s.

At the time the sample respondents in the fourth cohort of the SYPS reached the minimum school leaving age the government supported youth training programme was in a state of transition. The Youth Training Scheme (YTS) was being replaced by the substantially similar, but more flexible programme, Youth Training (YT). For the purposes of this analysis the programmes are treated as synonymous and the acronym YT is employed. It is important to emphasise, however, that the introduction of YT complicated evaluation practice in terms of defining what actually constitutes youth training. This is because one of the key elements of YT was increased flexibility, which meant that training could be made up of multiple occurrences, of variable length and source, over a greater period of eligibility. Moreover, many of these features are retained in current programmes.

For the purposes of our analysis YT participation is first identified from a survey question asking respondents if they had ever been on YT. The responses to this question were cross-checked with the diary information regarding the individual's main activity together with more general survey questions regarding YT, such as qualifications and training activity, as well as the number, length and type of courses attended. This allowed the construction of a variable labelled *YT Experience*, which records whether or not the respondent participated in YT after leaving school. Clearly, given the flexibility introduced with YT, and in order to more accurately capture the effect of the requisite training, training tenure also need to be taken into account. The survey gathered information from respondents on time spent in training with respect to their first or most recent YT course, dependent on the relevant sweep. Due to sporadic reporting on start and finish dates and a number of participants reporting participation in multiple courses, it was not possible to construct a reliable measure of training tenure directly from these responses. However, the survey also asked respondents the total time they had spent on YT in the form of a grouped question.[139] Comparison of the responses across all the relevant questions pertaining to duration then allowed the construction of three variables to control for training tenure. *YT One* represents less than 12 months in training; *YT Two* represents at least 12 months and up to 24 months in training while *YT Three* represents more than 24 months on the programme.

Three further variables were constructed in order to capture other elements of YT that may impact on the various outcome measures. The first, *YT Employer*, identifies if the respondent's current job is with the employer who sponsored the YT placement. The significance of controlling for this factor is highlighted in Main and Shelly (1990) and Whitfield and Bourlakis (1991). On the one hand, remaining with the same employer may have a positive impact on earnings because it would imply that specific human capital would continue to be drawn on resulting in higher productivity than with a different employer. On the other hand, it could also indicate that training is ongoing, possibly as part of a longer scheme, and hence have a negative impact on earnings (Main and Shelly, 1990, p. 502). A second variable, *YT Qualifications*, reflects the increased emphasis introduced with YT on vocational qualifications emphasised by, among others, Mealli, Pudney and Thomas (1995).

[139] The options were: less than 6 months, 6 months or more but less than 12 months, 12 months, more than 12 months but less than 24, and 24 months or more.

This variable was constructed using data from a survey question asking respondents if they gained qualifications as a consequence of YT participation, which was cross-checked with other survey questions regarding vocational qualifications.[140] The third variable, *YT 1993*, was constructed to control for the possibility that the respondent entered or had not completed their YT course within the final 12 months of the survey.

6.3: Estimated Impact of Youth Training on Earnings

Enhancing the earnings of participants was not a stated aim of YT however, as discussed in chapter 2, it does go someway to addressing the more fundamental question of whether or not the individual benefits from participation in this form of training. The results of such analysis naturally impact on the attractiveness of the scheme and its successor developments to future school leavers. Wages can also be viewed as a productivity measure and hence a proxy for how employers value the scheme, indeed if the programme is successful in enhancing skills we would expect an increase in wages. In general, however, existing research indicates that participation in government-assisted youth training programmes is associated with a negative impact on subsequent earnings. This is perhaps not unexpected given that most youth datasets only cover the labour market experiences of 16 to 19 years olds. As noted by a number of analysts, including Main and Shelly (1991), Breen (1992), Dolton et al., (1994 and 1995), the immediate post-compulsory schooling period is typically one of substantial human capital investment and many individuals will not have completed their training. Clearly, we cannot observe any of the long-term benefits of their training which may also result in the current observed earnings of trainees being lower than those of non-trainees, "...not because their expected

[140] A number of earlier studies (Main and Shelly, 1990; Dolton et al., 1994b) emphasised the importance of controlling for completion of the relevant training programme (course). However, their analyses were conducted on YTS, which had a fixed duration of one or two years depending on the version of the programme. In the case of YT, there is no fixed programme length and multiple training events are possible. Moreover, there would be a high degree of multicollinearity between such a variable and the variables *YT Employer* and *YT Qualification,* as both these presuppose completion of the relevant course. The number of courses was also introduced as a control in some of the wage specifications tested below but it was found to be statistically insignificant. It is however, retained for the employment equation.

earnings stream is lower, but because we only observe that part of it which happens to be lower" (Dolton, et al. 1995, p. 269). Moreover, as we can only observe earnings of those actually in employment at the time of the follow-up sweep of the survey, any estimates of earnings impact are also subject to the respondent being employed.

Furthermore, as observed by Main and Shelly (1990), such estimates cannot fully capture the non-pecuniary aspects gained from participation in the training and subsequent employment, such as self-confidence. Neither can they capture gains to society, such as declines in criminal activity as discussed in chapter 2 this would require a far more detailed dataset. Nevertheless, working within these limitations, analysis on the earnings impact of training does provide an initial appraisal of relative programme performance compared with alternate school-to-work transition routes based on a common metric and the opportunity to assess the effect of a variety of factors on the respondents' earning potential. The evaluation of youth training programmes in this fashion also allows us to assess the relative impact on the potential earnings of different 'types' of individual. This last attribute is particularly useful if the policy objectives of the programme specifically 'target' different groups, such as early or unqualified school leavers

This section reports the results of a sensitivity analysis on the estimated earnings of a sample of respondents drawn from the fourth cohort of the SYPS in order to determine the any variance in effect due to the adoption of different evaluation methodologies. This is preceded by a preliminary examination of the earnings data in the SYPS, the determination of an appropriate baseline specification and a discussion on the factors involved in the decision on whether or not to participate in youth training.

6.3.1: Earnings Data in the SYPS

The dependent variable used in the analysis is the log of gross hourly earnings and the sample comprises all those young people who had entered the labour market by October 1993, for whom the dependent variable was greater than zero, and complete records for the relevant variables were available. There were 1,293 survey respondents in the labour market sample of which 1,044 were in employment and 908 had valid reported earnings.

There were two questions in the survey that specifically relate to earnings:

> Last week, what was your pay from this job or scheme? (before deductions of tax, etc., please work out amount for one week)

And

> Adding together all the money you received, what was your total income last week? (include wages, benefits grants, please work out amount for one week, if nothing write '0').

Drawing on the first of these together with responses to a question asking respondents how many hours they worked in their job or scheme last week an hourly earnings variable was computed. Taking into account that the SYPS earnings data are self-reported, a series of cross checks was adopted in order to establish internal validity. At the recording stage the survey and coding instructions prompted checks for ambiguities and corrected for these where appropriate. For example, converting monthly earnings to weekly earnings where the respondent indicated this was the case. At the stage of computing the earnings variable, responses were compared with the second survey question and cross checks made with questions such as those regarding income from other sources (welfare, family, etc.), occupation and employment status. This allowed a number of more extreme figures to be adjusted or eliminated.[141]

The hourly earnings figure were then adjusted to a gross measure through reference to taxation and national insurance regulations in force in October 1993 while hours worked in excess of 39 were assumed to be overtime and paid at time and a half. The completion of these adjustments resulted in a measure of gross hourly earnings, or basic wage rate, for each respondent. Table 6.2a reports the summary statistics for the resultant SYPS earnings sample by YT status and gender. It is readily apparent

[141] For example, four respondents gave the answer 168 when asked: "How many hours did you work last week?" – clearly a case of the right answer to the wrong question. Additionally, a number of respondents with comparatively low earnings were subsequently identified as having been recorded as unemployed and their earnings were found to correspond with the rate of unemployment benefit in force at the time.

that the recorded wage rates display considerable variation, ranging from just 88 pence per hour to more than 16 pounds per hour. However, with respect to the lower rates, it should be noted that there was no minimum wage legislation in force at the time and the survey did not adequately account for 'payments in kind', such as board and lodgings.[142] Indeed many of the lower earnings were associated with occupations that are notoriously low paid, such as childcare and the hotel industry which, at the same time, are also associated with the provision of non-pecuniary benefits, like accommodation.[143] This may result in an underestimation of the wage rate in the case of some respondents. On the other hand, at the other end of the scale, the majority of high earnings were recorded by either the self-employed or participants in the oil and gas industry. The latter are associated with higher earnings, however some of the self-employed figures may be inflated if, for example, the respondent either failed to take into account tax and national insurance contributions and/or underestimated the number of hours they actually worked in the survey week. It is also possible that some self-employed respondents may not have adequately accounted for income received during, but earned prior, to the survey week. The external validity of the data is discussed below.

Table 6.2a: Reported earnings in SYPS sample by gender and YT status*

In Paid Employment Autumn 1993 (£/hr)		YT ($N = 317$)	Non-YT ($N = 591$)	Whole sample ($N = 908$)
Gross hourly wage rate	*Males*	3.94 (1.16)	4.54 (1.78)	4.35 (1.60)
	Females	3.85 (1.28)	3.99 (1.32)	3.95 (1.31)
	Average	3.89 (1.22)	4.23 (1.56)	4.11 (1.46)
	Range	1.29 – 8.21	0.88 – 16.88	0.88 – 16.88

* *Standard deviations in parentheses*

142 It is feasible that the recent introduction of minimum wage legislation in the UK will be accompanied by a watering down of such non-pecuniary benefits helping offset the cost of such legislation to employers. However, data detailing the receipt of non-pecuniary benefits before and after the introduction of the legislation would be required to accurately assess its impact on low wage employees.

143 Amongst the low earners there was also one respondent whose occupation was recorded as 'professional athlete'. Professional sports-people also often receive substantial payments in kind in the early stages of their career and their income can be extremely variable at the best of times.

Inspection of table 6.2a also provides an indication of the likely earnings impact of participation in YT. Such preliminary investigation suggests that participation in YT is associated with lower average earnings. This is particularly pronounced in the case of male youth, where the earnings of participants (YT) averaged just £3.94 per hour compared with £4.54 for non-participants (Non-YT). However, as discussed in section 3.4 (chapter 3 of this thesis) such simple comparisons do not take into account any differences in the labour market, background or family characteristics embodied by those on YT relative to those choosing other routes. This issue is addressed below where the impact of training on post-programme earnings is estimated using a range of more comprehensive models.

In table 6.2b the SYPS earnings data are compared with data from the New Earnings Survey (NES) as a means of establishing an approximate measure of external validity, or representativeness of the sample. The NES rates are national figures for 18 to 20 year old youth. Differences between the two sets of figures are significant when comparison is based on the (approximate) manual/non-manual distinction used by the NES. In the case of male youth, the comparison suggests that the SYPS earnings data tend to be higher than the population average particularly in the case of non-manual occupations (£4.53 compared to £4.37). The earnings of females in the SYPS earnings sample are again higher but variability is more substantial in the case of manual occupations (£3.60 compared to £3.46). These differences may be partly mitigated by the fact that the NES sample date is April 1993 compared with October 1993 for the SYPS.

Table 6.2b: Reported earnings in SYPS sample by gender and occupation: comparison with the 1993 New Earnings Survey (NES)*

Gross hourly wage rate (£/hr)	*Males*	*Females*	*Whole Sample*
NES:[1]			
Manual	4.11 (1.0)	3.46 (1.3)	3.79 (avg.).
Non-manual	4.37 (1.3)	4.08 (0.7)	4.22 (avg.).
All	**4.21 (0.8)**	**3.94 (0.6)**	**4.08 (avg.)**
SYPS:[2]			
Manual	4.18 (1.61)	3.60 (1.53)	3.96 (1.61)
Non-manual	4.53 (1.50)	4.12 (1.12)	4.22 (1.24)
All	**4.35 (1.60)**	**3.95 (1.31)**	**4.11 (1.46)**

Source: [1] *ONS, (1993) New Earnings Survey*
[2.] *SYPS earnings sample (see text)*
Notes: * *Standard deviations in parentheses*

Comparison with 1994 NES data did not, however, reveal substantive changes in the youth earnings other than in the case of females in manual occupations with the reported wage rate increasing to £3.63. In contrast, the rate for females in non-manual occupations recorded a marginal fall leaving the overall rate for young women unchanged. In the case of young men, the rate for those in manual occupations showed a marginal increase while the non-manual rate remained unchanged resulting in a slight increase in the overall rate to £4.23. The differences between the two sets of figures may also be due to the fact that no occupational classification was recorded for nearly one fifth of the SYPS respondents. On the other hand, if we focus on the figures for the groups as a whole (in **bold text**), it can be seen that the two sets of figures are fairly similar and hence the SYPS earnings data would appear to be reasonably representative of the target population.[144] Good practice would suggest, however, that a caveat stating that male earnings in the fourth cohort of the SYPS may be inflated should be appended to any analysis.

[144] Note that the NES does not provide statistics on the earnings for males and females combined and hence the figures used in column three (*Whole Sample*), in the upper portion of table 6.2b, are thus simple averages.

6.3.2: Choice of Baseline Specification

Drawing on human capital theory we know that earnings are related to a variety of personal and background characteristics such as gender, education and local labour market conditions.[145] Hence, in order to assess the effect of YT participation on earnings potential, it is necessary to adopt a more comprehensive estimation procedure capable of taking these characteristics into account. One approach is to use OLS to regress the basic hourly wage on those factors that are most likely to influence earnings, including the training variables defined in section 6.2. It is highly likely that the resulting estimates will be subject to sample selection bias, but this first generation approach serves as an informative starting point for our analysis in helping to determine the appropriate specification.

The SYPS contains data on a range of different characteristics allowing a number of specifications to be constructed and compared in order to determine a suitable baseline. To aid discussion the explanatory variables are grouped into five broadly related categories although the individual variables are broadly in accordance with the studies reviewed in chapter 4. The initial focus is on the respondent's *personal characteristics* that may impact on their earnings potential. It has already been observed above that young women generally earn less than their male counterparts. Women comprise more than half (54 percent) of the SYPS earnings sample and hence a dummy variable labelled 'Female' is included to control for gender.[146] A number of studies, including Dolton et al. (1995) and Green et al. (1996) suggest that marital status is also an important determinant of earning potential. Given that nearly 8 percent of the sample reported that they were married a variable is thus included to control for this factor. To the extent that marriage reflects stability and a willingness to undertake responsibility we might reasonably expect the variable to be associated with higher earning potential.

---

[145] Although the dominant paradigm in modern labour economics, human capital theory is primarily used here in order to elucidate the modelling strategy. Other interpretations of the processes involved are equally plausible. For example, Main and Shelly (1991) use the institutional construct of a labour queue proposed by Thurow (1975) in his 'job competition model' in the sense that better jobs are awarded to those at the front of the queue.

[146] All dummy variables take the value 1 if the characteristic is present or 0 otherwise.

The next group of variables relates to the individual's formal education or *school experience*. The number of examinations passed is the usual measure of an individual's educational achievement and represents accumulation of human capital. Given employers often equate exam success with ability, and therefore potential productivity, we would expect higher educational achievement to be associated with higher earning potential. Table 6.3 hence reports the ordinary hourly wage rates computed from the SYPS earnings sample summarised by the level of school leaving qualification and YT status. Five different categories are used to summarise the examination data in the SYPS ranging from unqualified (no passes awarded) to those gaining Higher awards. The figures in the first column of the table (*in Italics*) report the composition of the sample in terms of the proportion of young people attaining the respective levels of school leaving qualification.

It can be seen in table 6.3 that the expected earnings pattern is evident in the case of YT participants and holds for the sample as a whole excepting in the case of those who left school with Higher passes. With respect to non-participants the pattern is more mixed with earnings of the two extremes of the qualifications spectrum (Unqualified and Highers) appearing to be out of line with expectations. This may in part be due to differences in labour market experience. For example, many young people who remained at school to obtain Higher passes would have as little as 12 months in the labour market compared with approximately 36 months for those who left school at the earliest possible opportunity. Again, the general impression gained from the table is that participation in YT is associated with lower earnings than those received by non-participants. The class omitted in the regression analysis is 'O/S levels with D and E grades only'.

Table 6.3: SYPS earnings by school leaving qualification*

In Paid Employment 1993 (£/hr)	*Sample mean*	YT ($N = 317$)	Non-YT ($N = 591$)	Whole Sample ($N = 908$)
One or more Highers	*0.34*	4.16 (1.28)	4.22 (1.28)	4.22 (1.28)
Four or more O/S-levels	*0.19*	4.02 (1.17)	4.45 (2.12)	4.31 (1.87)
One to three O/S-levels	*0.24*	3.75 (1.15)	4.24 (1.52)	4.03 (1.39)
O/S-levels with D & E grades only	*0.14*	3.94 (1.35)	3.89 (1.22)	3.92 (1.29)
Unqualified	*0.09*	3.70 (1.11)	4.03 (1.87)	3.84 (1.49)
Average		3.89 (1.22)	4.23 (1.56)	4.11 (1.46)

** Standard deviations in parentheses*

An alternative approach to account for formal education performance is to use a composite variable such as the ILEA integer index used by O'Higgins (1994) and Payne (1995) or EDUCAT as used by Dolton et al. (1994b) among others. The equivalent variable in the SYPS dataset is TOTSCEP, which is an integer index taking the values of 1 to 14 dependent on different increments of achievement in school examinations. These range from no awards at O/S levels (or grades of 6 or 7 in S levels) through to 6 plus Higher passes.[147] The average level of overall attainment at school according to the TOTSCEP index for the SYPS earnings sample was 6.10, with a corresponding standard deviation of 3.70. Both these approaches are used in the initial analysis reported below.

A further two schooling experience variables are included, S4 Leaver to capture the effect on earnings of leaving school at the earliest opportunity (approximately 44 percent of the sample), and S5 Leaver to capture the effect of remaining at school for a further year (37 percent). Both variables also approximate the individual's school leaving age (16 and 17 years respectively) and reflect their relative labour market experience (3 or 2 years). On the one hand, leaving school early may have a negative impact on earnings because the individual would have had less opportunity to accumulate human capital. On the other hand, if they entered directly into

[147] A breakdown of the TOTSCEP variable is provided in the bottom half of table A3, in the appendix to this thesis.

employment, as approximately three-quarters of the S4 leavers did, their earnings may benefit from the accumulation of specific human capital gained through experience on the job.

Labour market experience is detailed in the following category both in the current job and more generally. General demand conditions are included in the form of the 'Local unemployment rate' variable, which reflects localised job opportunity. This variable gives the local unemployment rate as measured for the school-travel-to-work area (ONS data) and rates ranged from 2.8% in Aberdeen to 16.9% in Cumnock and Sanquhar. The average unemployment rate recorded in the earnings sample using this measure was 9.06 percent (3.61) compared with national (Scottish) unemployment rate at the time of 9.5 percent. The survey also gathered information from respondents on time spent in employment with respect to their first or most recent job dependent on the relevant sweep. Again, due to sporadic reporting on start and finish dates and a number of participants reporting multiple jobs, it was not possible to construct a reliable measure of employment tenure directly from these responses. However, the survey did ask respondents about their experience of unemployment and after cross-checking these responses with the diary and school leaving information it was possible to construct a variable measuring unemployment duration in months. The value of this variable ranged from no unemployment to as much as 36 months (that is, no employment over the duration of the survey), with an average of just under 2 months (3.92). We would expect that prolonged duration of unemployment would be associated with a negative impact on earnings.[148]

The number of hours worked in the sample week is included to control for those respondent's who reported working overtime or less than full-time in the survey week. On average respondents' worked for approximately 38 hours (7.22) in the sample week. Two further aspects of the respondent's job are also controlled for in the base model. The first is a dummy variable indicating whether or not the individual reported receiving training in their current job. The survey asked a range of questions about current training status including details on delivery, level and

---

[148] The corresponding standard deviations are given in parentheses.

occurrence.[149] Nearly all respondents gave an affirmative response to a question asking them if they received training in their current job but investigation of the more detailed questions revealed that much of this training was received from 'work-mates'. It was decided to only include training that the respondent indicated as being received from a supervisor, involving a course or as leading to a qualification. In other words, the 'Does training' variable indicates the receipt of 'formal training' in the respondent's current job. Approximately 25 percent of the earnings sample reported receiving training using this definition. It was also decided to include a variable indicating if the individual was self employed given the observation made above, in the dependent variable description, pertaining to the higher earnings that were associated with this form of work in the sample.

Industry and occupation variables are included in the next category to account for some of the variation due to differences in prevailing wage rates across different sectors and occupational choice. A case can be made for either the inclusion or exclusion of these variables. For example, Whitfield and Bourlakis (1991, p.50) put forward the argument that government assisted programmes channel participants into lower paying sectors where placements are concentrated. They hence argue that industry and occupation variables should be included to allow for the possibility that trainees may not be randomly distributed across sectors. On the other hand, Main and Shelly (1990, p. 503) suggest that, if part of the effect of the programme is to influence the type of job held the inclusion of descriptors might reduce the measured YT effect. Equally, they stress that the industrial and occupational descriptors may capture important non-pecuniary effects and, hence, should be included (op cit., p. 503). Here a similar approach is followed to that adopted by Main and Shelly in estimating the wage equations with and without this category. The final category of variables relates to participation and the nature of the individual's YT experience. The relevant variables were described above in section 6.2.

---

[149] Dolton et al., (1995) used responses to similar questions in the YCS (England and Wales) dataset to investigate comparative returns to training provision in the public and private sectors. By drawing on the detail in the different questions they were also able to construct different training profiles and estimate the effect on subsequent earnings of choosing different forms of training within the context of a polychotomous choice model.

Three specifications of the wage equation were constructed using these variables and the resultant OLS estimates are reported below in table 6.4. The first specification was run without either the industrial/occupational descriptors or the variables controlling for time spent on the YT programme. The industrial and occupational descriptors are then introduced in the second specification followed by the addition of controls for training tenure in the third.

In each of the specifications the results largely conform to our prior expectations. In particular, superior educational attainment is clearly associated with higher subsequent earnings thus ironing out the apparent contradictions noted when describing the data. This relationship was reinforced when the regressions were run using the alternative education measure TOTSCEP the resultant coefficients were positive and significant, ranging from 0.014 (3.2) to 0.019 (4.68), depending on the specification employed. The impact of using this alternate measure on the remainder of the coefficients and model fit was minimal and given the ease of interpretation associated with the more detailed examination descriptors they are retained for the remainder of the analysis.[150] The OLS estimates also suggest that the extra experience gained from leaving school after S4 more than outweighs any negative impact that might be associated with comparatively less schooling. Whereas an extra years schooling does not appear to have as large an impact on potential earnings, albeit still positive and significant. However, It should be noted that all these estimates are related only to *early returns* in the respondents' life cycle of earnings.

---

[150] For example, when TOTSCEP was used in place of the examination descriptors in the estimation of specification three, the recorded coefficient and (*t-stat*) for *YT experience* was minus 0.120 (2.32) and the R^2 was 0.18.

Table 6.4: OLS wage equations for school leavers in employment, Autumn 1993

Descriptive variables	*Specification 1*	*Specification 2*	*Specification 3*
Personal characteristics			
Female	-0.118 (5.17)	-0.139 (5.68)	-0.141 (5.78)
Marital status	0.028 (0.69)	0.040 (0.98)	0.036 (0.89)
School experience			
Unqualified	-0.043 (0.95)	-0.029 (0.65)	-0.023 (0.51)
1-3 O/S-Grades	0.061 (1.65)	0.039 (1.09)	0.047 (1.30)
4+ O/S-Grades	0.091 (2.32)	0.048 (1.24)	0.053 (1.35)
Highers	0.144 (3.32)	0.095 (2.16)	0.099 (2.24)
S4 Leaver	0.126 (2.88)	0.123 (2.86)	0.129 (3.01)
S5 Leaver	0.084 (2.44)	0.079 (2.33)	0.077 (2.31)
Labour market experience			
Local unemployment rate	-0.004 (1.40)	-0.005 (1.76)	-0.005 (1.80)
Unemployment duration (mths)	-0.008 (3.02)	-0.008 (2.84)	-0.008 (2.94)
Hours worked	-0.010 (7.27)	-0.010 (7.64)	-0.010 (7.63)
Does Training	0.063 (2.47)	0.047 (1.86)	0.048 (1.89)
Self Employed	0.177 (1.95)	0.195 (2.06)	0.192 (2.04)
Industrial and occupational descriptors			
Mining and utilities	..	0.194 (1.71)	0.192 (1.70)
Manufacturing	..	0.052 (0.64)	0.051 (0.63)
Construction	..	-0.077 (0.92)	-0.085 (1.02)
Services	..	-0.075 (0.95)	-0.074 (0.95)
No industry given	..	-0.230 (1.55)	-0.212 (1.44)
Unskilled	..	-0.145 (4.67)	-0.143 (4.63)
Skilled Manual	..	-0.108 (3.19)	-0.104 (3.06)
Professional/Intermediate occupation	..	0.021 (0.51)	0.022 (0.54)
No occupation given	..	0.164 (2.01)	0.153 (1.87)
Youth Training			
YT Experience	-0.029 (0.45)	-0.044 (1.17)	-0.122 (2.36)
YT One (< 12 months)	..	..	0.114 (2.41)
YT Three (> 24 months)	..	..	-0.038 (0.81)
YT Employer	-0.120 (3.21)	-0.108 (2.94)	-0.092 (2.48)
YT Qualification	0.023 (0.55)	0.032 (0.78)	0.084 (1.91)
Still on YT 1993	-0.046 (0.88)	-0.054 (1.04)	-0.003 (0.054)
Constant	6.294 (81.07)	6.484 (56.94)	6.479 (57.06)
F- statistic	7.24	7.09	6.99
Mean of dependent variable	5.9627	5.9627	5.9627
Standard deviation of dependent variable	0.339	0.339	0.339
Standard error of regression	0.322	0.314	0.312
R - squared	0.12	0.17	0.18
N	*908*	*908*	*908*

**coefficient (t-statistic)*

One variable that does not conform to expectations is the one representing training in the respondent's current job, which has a positive coefficient whereas *a priori* reasoning based on human capital theory, suggests that this variable should act in the opposite direction. The theory suggests that workers contribute to the cost of firm-specific training in the form of a lower wage during the training period, than would otherwise be the case, with the incentive of a higher wage subsequent to completion of the training (Becker, 1962).[151] The presence of some form of implicit contract that spreads the cost of the training over an extended period could be a possible explanation but suspect data is equally plausible. The survey asked respondents to indicate what form of training they received at work but they were not restricted to one option and many gave multiple responses. A later question asked them to indicate which from of training was the 'main' one but this still left a small number of inconsistencies that could have affected the estimates.

Conversely, we would hope that completed youth training events result in a positive impact, but the *YT experience* variable is associated with a negative impact on wages in all three specifications ranging from minus 3 to minus 12 percent. The larger of these, derived from the third specification, is the only estimate to reach conventional levels of statistically significance. The variables indicating if a participant remained on YT for an extended period (more than 24 months), had recent YT involvement, and/or remained in the employment of the YT sponsor are all associated with a further negative impact on potential earnings. On the other hand, it can be seen that these negative influences are partially offset if the trainee gained a qualification from the programme and/or their involvement was relatively brief (less than 12 months). In the YT case, however, the negative impact of training is consistent with previous estimates of the early returns to government assisted youth training programmes. This could be a reflection of a variety of factors ranging from lower wage

[151] Firm-specific training is hypothesised to increase productivity in the firm providing the training as opposed to general training, which is associated with increased productivity in more than one firm. Because of the existence of a bilateral monopoly in the labour market, human capital theory suggests the cost of firm-specific training will be shared by both workers and firms (see also Oi, 1962).

expectations among trainees to a greater supply of relatively skilled youth labour as discussed in chapters 3 and 4 (see also Hutchinson and Church, 1989).

In order to assess the impact of participation in YT terms of the trainees' wage rate it is necessary to focus on specific individual profiles. To facilitate discussion and allow ready comparability with the selection models presented below a stereotypical 'average' individual is hence defined as having left school on completion of S4 with one to three O/S levels. They reside in an area with an unemployment rate of 9 percent; have experienced 2 months unemployment (between April 1991 and October 1993) and work for 38 hours per week in the service sector in a skilled-manual occupation. In addition to these characteristics, the average YT participant spent between 12 and 24 months on the programme and obtained a qualification from that involvement. Using these definitions, together with the coefficients from column three of table 6.3, the wage rate for the average male participant is estimated at £3.93 per hour compared with the non-participation rate of £4.08. The corresponding wage rates for female youth are £3.41 and £3.54. In other words, according to the OLS equations, participation in YT is associated with a negative earnings impact in the order of minus 3.7 percent in the case of both the average female trainees. The wage rates of those respondents remaining with the employer who sponsored their YT training fared even worse earning up to 35 pence per hour less than the average participant.

As observed above, these estimates are however likely to be biased because orthodox regression models effectively assume random assignment to YT. That is, they do not take into account any unobserved differences between those that choose to participate in, or are assigned to, YT and those who remain without YT. In other words, the OLS model does not control for *sample selection bias*. Selection models are designed to account for unobserved differences and are examined in the next two sections.

6.3.3: The Participation Decision

The prototypical approach to estimating the effect of programme participation in the likely presence of sample selection bias requires that an attempt be made to model the participation decision. Given we do not directly observe the decision to participate, only the realisation of that decision, the process is typically approximated using a Probit model to estimate the probability of a school leaver having had experience of YT by the time the follow-up sweep is conducted. The model is constructed using variables that might help determine whether or not the young person participates in such training programmes while the data is drawn from the responses recorded in the first sweep of the survey.[152] The primary focus is on the individual's educational performance and their socio-economic background at the time they reached the minimum school-leaving age. The measures of educational achievement are restricted to O-levels, which would have been the highest grade attainable in S4, with the omitted class again being 'O/S levels with D and E grades only'. Clearly, educational performance impacts not only on the respondent's choice of school-to-work transition route but also on their chances of gaining a place in higher education and their acceptability to an employer, although it is not possible to separate these effects. Attendance of an independent school might be expected to operate in much the same way (Green, et al., 1996, p. 13). The recorded occupational status of the respondent's father is used to reflect socio-economic background, with the omitted class being 'unskilled-manual'. The father's employment status in April 1991 is also included as a further indicator of the young person's background and opportunity.

A variable indicating whether the individual habitually played truant while at school is included in an attempt to represent their attitude toward school. The survey asked respondents to indicate if they played truant and, if they did, how often. The truancy variable includes those who responded that they did so for 'days or weeks at a time'. The variable representing gender is again included to control for the different experiences of young men and women in the transition process. Finally the local unemployment rate, as recorded for the respondent's school travel-to-work area, is

---

[152] As such the participation Probit uses the whole sample, or 1,293 case, that is all those in the labour market at the time of the follow-up sweep (see figure 6.1). The descriptive statistics for this sample, which is also used to determine the impact of training participation on the probability of subsequent employment, are presented below in section 6.4, table 6.12. The Probit model was formally presented in section 3.5.2, equations 3.21 to 3.24.

included to reflect the possibility that young school leavers are effectively forced into YT due to limited job opportunities or little possibility of welfare. The results of this exercise are presented in table 6.5. As in the remainder of this chapter the estimates are derived using the econometric package LIMDEP 7.0.

Table 6.5: Probability that a school leaver has experience of YT*

Descriptive variables	*N=1293*
Female	0.123 (1.66)
Unqualified	0.466 (3.88)
1-3 O/S-levels	0.369 (4.04)
4+ O/S-levels	0.137 (1.30)
Truant (habitual)	0.287 (2.15)
Independent School	-0.609 (1.50)
Father unemployed	0.014 (0.91)
Father in skilled/ manual occupation	-0.357 (3.48)
Father in professional/ intermediate occupation	-0.714 (5.92)
Father's occupation not given	-0.260 (2.11)
Local unemployment rate	0.054 (5.12)
Constant	-0.605 (4.18)
Chi-square	124.81
Log-likelihood	-794.07

**coefficient (t-statistic)*

The estimates in the table show participation in YT to be most directly influenced by the level of attainment at school, the respondent's attitude to school and local unemployment conditions. On the other hand, those respondents with a father in a professional/intermediate or skilled/manual occupation and those with relatively better school leaving qualifications are less likely to participate. Otherwise there seems to be a relatively weak connection between the remaining personal characteristics and participation in YT.

Heckman (1979) illustrated that the estimates in table 6.5 can subsequently be used to correct for probable sample selection bias by treating it as an omitted variable and thus allowing the estimation of the 'corrected' wage equation(s). The Heckman two-step model was detailed formally in chapter 3 (section 3.5.2, equations 3.7a to 3.11b) and the results of applying this widely used method are reported at the beginning of the next section. Before proceeding however, it is necessary to address two often-cited concerns regarding the use of a Probit in order to model the selection process, namely non-normality and heteroskedasticity (Breen, 1996, p. 70).

In order to test the relevant hypotheses we utilise likelihood ratio tests following Orme (1995). These tests are similar in spirit to the auxiliary regression tests of heteroskedasticity (White, 1980) or serial correlation (Breusch and Pagan, 1979) in which the significance of an extra set of regressors is tested by means of asymptotically valid likelihood ratio test statistics.[153] To test for non-normality the Probit, $y_i^* = X_i\beta + \varepsilon_i$,[154] is first estimated in the standard fashion in order to obtain maximum likelihood estimates of the parameter coefficients, $\tilde{\beta}$, and the maximised log-likelihood, $LogL_0$. Squared and cubed terms of the parameter estimates, $(X_i\tilde{\beta})^2$ and $(X_i\tilde{\beta})^3$, are then derived and an auxiliary regression is run including these in order to derive a new maximum log-likelihood, $LogL_N$. The appropriate test statistic is then $2(LogL_N - LogL_0)$ which should be distributed as a $\chi^2(2)$ under the null of normality. In this case the derived test statistic was 4.45 compared with the five-percent critical $\chi^2(2)$ value of 5.99, hence we cannot reject the null of normality.

Similarly, to test for heteroskedasticity the Probit is first estimated in order to obtain maximum likelihood estimates of the parameter coefficients, $\tilde{\beta}$, and the maximised log-likelihood, $LogL_0$. Test variables of the form $(X_i\tilde{\beta}).z_i$ are then included in an auxiliary regression, where z_i represents an m-vector of variables, which may cause heteroskedasticity, in order to derive a new maximum log-likelihood, $LogL_H$. The appropriate test statistic is then $2(LogL_H - LogL_0)$ which should be distributed as a $\chi^2(m)$ under the null of homoskedasticity. In this case the auxiliary regression was

153 Alternatively scores test may be used. These involve the derivation of higher moments of the appropriate distribution, which is fairly involved in the limited dependent variable case. Studies by Bera, Jarque and Lee (1984) and Chesher and Irish (1987), however, define the necessary recursions required to permit the sequential derivation of conditional expectations without recourse to integration.

154 Where y^* is an unobserved latent variable that determines the value of the dichotomous outcome variable, y (in section 3.5.2, $y^* = 1$ was used to denote participation and $y^* = 0$ non-participation, see equations 3.21 to 3.24).

run with all of the variables (omitting the intercept) and the derived test statistic was 13.74 compared with the five-percent critical $\chi^2(11)$ value of 19.67, hence we cannot reject the null of homoskedasticity.[155]

Section 6.3.4 presents the estimates of earnings impact derived from a range of commonly employed selectivity models using the participation Probit derived above. The models are presented in the same order as was adopted in section 3.5, which detailed the underlying theoretical constructs of the different methods. The resultant impact estimates, in terms of comparing the predicted wage rates of an average school leaver dependent on whether or not they participate in YT, are summarised at the end of the section in table 6.12, which also details the impact of different forms of YT experience.[156]

6.3.4: Earnings Estimates

Table 6.6 reports the estimated earnings equation following Heckman's two-step model, or limited information maximum likelihood estimation, which utilises OLS in the second step. The first column records the parameter estimates for a randomly chosen school-leaver should they participate in YT and the second column if they do not. With respect to the direction of impact, the different variables generally have the same signs as the OLS estimates that were reported in table 6.4. There are some differences between the participation and non-participation equations observed in the signs of the different industrial/occupational descriptors but these are likely to be due to the large proportion of participants who did not supply responses to the occupation and/or industry questions.[157]

On the other hand, the estimated earnings impact, in terms of our average individual (as previously defined), is substantially different to that of the OLS model. A male respondent with an average set of characteristics would have earned £3.77 per hour if

---

[155] LIMDEP incorporates the heteroskedastic Probit model in its 'canned' procedures and the Maxit command can be appended to derive the test statistics.

[156] The requisite equation numbers from chapter 3 (section 3.5) are given for each model immediately after it is introduced.

[157] The resultant relative lack of variation in some of these descriptors limits the application of more sophisticated models, such as the Multinomial Logit model with selection; see discussion below.

rewarded according to the participation equation whereas they would have earned £3.98 per hour if rewarded according to the non-participation equation. That is, an earnings impact associated with participation in YT in the order of minus 5.3 percent compared to minus 3.7 percent in the OLS case. In contrast, the corresponding rates for female youth were £3.43 and £3.42 per hour a marginally positive impact again compared with approximately minus 3.7 percent predicted in the OLS equation.

From the table it can be seen that the sample selection term (inverse Mills ratio or Lambda) does not reach conventional levels of statistical significance in either equation. This is generally taken as an indicator that sample selection bias is not a serious factor in the data (Heckman, 1979; and Melino, 1982) and that the impact can be derived using conventional OLS techniques, see for example Whitfield and Bourlakis (1994, p. 52). In any case, the two-step model has been roundly criticised in a number of recent investigations into the relative performance and the necessary (restrictive) conditions dictating efficient estimation when using this method.[158] This has led a number of practitioners to suggest that the two-step model only be used to test for the presence of sample selection and/or to obtain the starting values for estimating the model using full information maximum likelihood estimation (MLE). See, for example, Davidson and MacKinnnon (1993, p. 545), Greene (1993, p. 642) and see also the comparative MLE estimates below and the discussion in Breen (1996, pp. 69-72).

[158] For example, Leung and Yu (1996) conducted Monte Carlo experiments to help determine what conditions are necessary for the two-step model to produce robust estimates. Similar studies have been conducted by, among others, Stolzenberg and Relles (1997), Nawata (1994) and Paarsch (1984), these were reviewed in chapter 3 (section 3.6.2).

Table 6.6: Probit/OLS (Heckman Model) wage equations for school leavers in employment, Autumn 1993

Descriptive variables	*YT*	*Non-YT*
Personal characteristics		
Female	-0.094 (2.42)	-0.153 (4.90)
Marital status	0.093 (1.57)	-0.002 (0.04)
School experience		
Unqualified	-0.026 (0.48)	-0.066 (0.84)
1-3 O/S-Grades	0.021 (0.42)	0.057 (0.99)
4+ O/S-Grades	0.053 (1.01)	0.074 (1.31)
Highers	0.107 (1.53)	0.126 (2.12)
S4 Leaver	0.165 (1.52)	0.124 (2.48)
S5 Leaver	0.126 (1.23)	0.071 (1.93)
Labour market experience		
Local unemployment rate	-0.002 (0.33)	-0.012 (2.56)
Unemployment duration (mths)	-0.007 (1.72)	-0.008 (2.43)
Hours worked	-0.012 (4.86)	-0.010 (6.17)
Does Training	0.076 (2.03)	0.026 (0.79)
Self Employed	0.021 (0.17)	0.034 (2.64)
Industrial and occupational descriptors		
Mining and utilities	-0.049 (0.16)	0.220 (1.66)
Manufacturing	0.084 (0.71)	0.065 (0.61)
Construction	0.027 (0.22)	-0.109 (1.00)
Services	-0.064 (0.55)	-0.041 (0.40)
No industry given	-0.546 (1.70)	-0.108 (0.62)
Unskilled	-0.057 (1.11)	-0.176 (4.64)
Skilled Manual	-0.108 (2.14)	-0.076 (1.72)
Professional/Intermediate occupation	0.163 (2.31)	-0.031 (0.64)
No occupation given	0.436 (2.96)	0.056 (0.58)
Youth Training		
YT One (< 12 months)	0.097 (2.26)	..
YT Three (> 24 months)	-0.025 (0.57)	..
YT Employer	-0.082 (2.19)	..
YT Qualification	0.073 (1.83)	..
Still on YT 1993	-0.037 (0.71)	..
Lambda	-0.070 (0.70)	-0.136 (1.28)
Constant	6.330 (26.94)	6.442 (42.48)
F- statistic	3.28	5.41
Mean of dependent variable	5.9153	5.9881
Standard deviation of dependent variable	0.317	0.349
Standard error of regression	0.275	0.316
N	*317*	*591*

**coefficient (t-statistic)*

Little (1985) has also expressed reservations concerning the robustness of the type of two stage sample selection correction being used here, on the basis that identification rests on the inclusion in the YT participation equation of variables that are good predictors of YT participation but have no association with earning power. If this requirement cannot be met, then identification rests in the non-linearity of the Probit equation. Here we have chiefly used variables indicating the occupational status of the respondent's father as instruments that have an impact on the school-to-work transition route (including YT) but not necessarily influencing the wage. Thus the model is resting on both 'behavioural' and parametric assumptions. The instruments used, however, are consistent with those employed in most of the UK studies reviewed in chapter 4 and are also retained for all of the models subsequently estimated in this section. Nevertheless, the appearance of the remaining participation variables in the outcome equation could lead to collinearity problems, which would also limit the power of the *t*-test for sample selectivity on the coefficient of Lambda. The method proposed by Puhani (2000), is hence used to test for collinearity. This test, which builds on the work of Leung and Yu (1996), rests on the calculation of the 'condition number' for the regressors.[159] If the condition number exceeds 20 an alternative model, such as OLS, is likely to lead to more robust estimates. In this case the relevant condition number was 14.28 suggesting that estimates derived from a selectivity model (two-step or MLE) would be relatively more robust. A simpler technique is proposed by Breen (1996, p. 70) who suggests deriving the Lambda term from the Probit and then regressing this term against the explanatory variables being used in the wage equation. If the resultant R^2 is close to one he suggests using MLE techniques, if it is close to zero he suggests using OLS. In this case the R^2 was 0.91 indicating that MLE would be the preferred technique.

The preference for maximum likelihood estimation is a common theme in recent literature on sample selection bias, because a significant shortcoming of the Heckman two-step method arises from the use of OLS in the second step, which results in estimates that are consistent but not fully efficient. MLE estimates for the

---

[159] The necessary LIMDEP code to carry out this test is presented by Puhani in the appendix to his article.

outcome (wage) equation, which, on the other hand, are efficient, may now be obtained within LIMDEP (and other programmes) using incorporated routines. Although the focus of the analysis in this section is on the different extensions to the widely employed Heckman model that have been utilised in the literature, the comparative MLE estimates were derived to allow some means of validating the two-step estimates. The resultant wage rates for male youth, derived using the method of MLE were estimated at £3.83 per hour for participants and £4.14 per hour for non-participants. The corresponding wage rates for female youth were estimated at £3.49 and £3.50 per hour. The nominal rates are clearly higher in both cases but, in terms of the net impact on the wage rate associated with participation in YT, the divergence between the two methods is only reflected in the case of male youth. The estimated impact for males was approximately minus 7.5 percent when MLE were used compared to minus 5.3 percent when the two-step method was used. In contrast, the estimated earnings impact derived from the two methods for female youth is more closely aligned, being marginally negative when MLE is used as opposed to marginally positive in the two-step case. The divergence in the male case may, in part, be due to potential problems in the SYPS data on male earnings (non-representativeness) identified in section 6.3.2. Nevertheless, it would appear that the two-step estimates are not greatly out of line with the MLE corrections in this case. The full results for the MLE estimates of the Heckman model are provided in the appendix, table A 5.[160]

In spite of all these reservations use of the Heckman procedure has proliferated widely in the evaluation literature and, to a great extent, it still remains the industry standard. It is also the most widely applied model in existing evaluation of youth training interventions in the UK and, given one of the aims of this investigation is to explore the effect of using the two-step model relative to alternative models applied to the same data-set, it is a useful starting point.

---

[160] As previously noted the major advantage of the two-step method is its relative ease of computation. This is reflected in the number of iterations required to reach convergence in the different techniques, for example in this case MLE required 52 iterations for the participation equation and 47 for the non-participation equation in this compared with only 5 (for the Probit) in the two-step method. Clearly with improved computing power this is less of an impediment to MLE estimation.

The next two tables (6.7 and 6.8) present the results arising from the application of the Treatment Effects and Logit/OLS models (equations 3.12a to 3.12b and 3.13a to 3.16). The treatment effects model (Barnow et al., 1981) is very similar to the Heckman two-step model as it still rests on the Probit selection equation to model the participation decision. In contrast to the Heckman approach, this model only employs a single wage equation drawing on all the available data (908 cases) and includes *YT experience* as an additional descriptive variable. The advantage of this model is that it allows the effect of training to be more readily observed, as the coefficient of the YT experience dummy can be interpreted in the same fashion as in a conventional regression. The requisite value is minus 2.3 percent and statistically insignificant compared with minus 12.2 percent (and significant) when the wage equation was estimated using OLS. Our current focus, however, remains on the effect of training on the wage rate of an average respondent. In this case, the average male would have earned £4.15 per hour from participation in YT compared to a non-participant rate of £3.91, a positive difference of 5.7 percent. The positive impact is maintained for female youth with the corresponding rates of £3.62 and £3.40 per hour, or approximately 6 percent due to participation on YT for 12 to 24 months and obtaining a qualification. As noted above, YT experience on its own is however still associated with a negative impact on earnings, albeit small and statistically insignificant, as are remaining on the programme for an extended period or remaining with the YT employer.

Lee (1983) described a reformulation of the selection model allowing for more general specifications of the two-step approach. The most common of these is the Logit/OLS model (equations 3.13a to 3.16). It uses a Logit (Logistic regression) to model the participation decision and, as might be expected, the results are very similar to those from the Probit/OLS.[161]

---

[161] The logistic distribution has a variation of $\pi^{2/3}$, and consequently the β estimates obtained from the Logit have to be multiplied by $3^{1/2}/\pi$ to be comparable to those estimates obtained from the Probit. Amemiya (1981) demonstrated that the Probit and Logit parameter estimates would generally differ in magnitude according to relatively fixed proportions and found that multiplying the Logit estimates by 0.625 produces a close approximation to the distribution function of the standard normal.

Table 6.7: Treatment Effects Model wage equations for school leavers in employment, Autumn 1993

Descriptive variables	*Combined sample*
Personal characteristics	
Female	-0.139 (5.72)
Marital status	0.037 (0.94)
School experience	
Unqualified	-0.048 (0.80)
1-3 O/S-Grades	0.041 (0.84)
4+ O/S-Grades	0.039 (1.25)
Highers	0.100 (2.31)
S4 Leaver	0.129 (3.06)
S5 Leaver	0.077 (2.34)
Labour market experience	
Local unemployment rate	-0.007 (1.78)
Unemployment duration (mths)	-0.008 (3.05)
Hours worked	-0.010 (7.74)
Does Training	0.049 (1.95)
Self Employed	0.195 (2.10)
Industrial and occupational descriptors	
Mining and utilities	0.191 (1.72)
Manufacturing	0.049 (0.61)
Construction	-0.086 (1.05)
Services	-0.075 (1.00)
No industry given	-0.214 (1.47)
Unskilled	-0.143 (4.71)
Skilled Manual	-0.102 (3.06)
Professional/Intermediate occupation	0.022 (0.54)
No occupation given	0.152 (1.90)
Youth Training	
YT experience	-0.023 (0.14)
YT One (< 12 months)	0.112 (2.41)
YT Three (> 24 months)	-0.038 (0.83)
YT Employer	-0.092 (2.54)
YT Qualification	0.085 (1.96)
Still on YT 1993	-0.005 (0.94)
Lambda	-0.061 (0.62)
Constant	6.463 (56.44)
F- statistic	6.75
Mean of dependent variable	5.9627
Standard deviation of dependent variable	0.339
Standard error of regression	0.307
N	*908*

**coefficient (t-statistic)*

Table 6.8: Logit/OLS (Lee Model) wage equations for school leavers in employment, Autumn 1993

Descriptive variables	*YT*	*Non-YT*
Personal characteristics		
Female	-0.093 (2.42)	-0.153 (4.90)
Marital status	0.093 (1.57)	-0.002 (0.04)
School experience		
Unqualified	-0.031 (0.49)	-0.067 (0.84)
1-3 O/S-Grades	0.021 (0.41)	0.057 (0.98)
4+ O/S-Grades	0.054 (1.02)	0.075 (1.32)
Highers	0.108 (1.54)	0.126 (2.12)
S4 Leaver	0.165 (1.51)	0.124 (2.48)
S5 Leaver	0.126 (1.22)	0.071 (1.94)
Labour market experience		
Local unemployment rate	-0.002 (0.35)	-0.013 (2.58)
Unemployment duration (mths)	-0.007 (1.73)	-0.009 (2.41)
Hours worked	-0.012 (4.86)	-0.010 (6.16)
Does Training	0.076 (2.03)	0.026 (0.79)
Self Employed	0.021 (0.18)	0.344 (2.65)
Industrial and occupational descriptors		
Mining and utilities	-0.048 (0.16)	0.220 (1.66)
Manufacturing	0.084 (0.71)	0.065 (0.61)
Construction	0.027 (0.22)	-0.109 (1.00)
Services	-0.064 (0.55)	-0.041 (0.40)
No industry given	-0.546 (1.70)	-0.108 (0.62)
Unskilled	-0.057 (1.18)	-0.176 (4.65)
Skilled Manual	-0.108 (2.14)	-0.076 (1.72)
Professional/Intermediate occupation	0.162 (2.31)	-0.031 (0.64)
No occupation given	0.435 (2.96)	0.059 (0.58)
Youth Training		
YT One (< 12 months)	0.097 (2.26)	..
YT Three (> 24 months)	-0.024 (0.57)	..
YT Employer	-0.082 (2.20)	..
YT Qualification	0.074 (1.83)	..
Still on YT 1993	-0.037 (0.71)	..
Lambda	-0.073 (0.73)	-0.138 (1.30)
Constant	6.334 (26.95)	6.440 (42.50)
F- statistic	3.28	5.41
Mean of dependent variable	5.9153	5.9881
Standard deviation of dependent variable	0.317	0.349
Standard error of regression	0.275	0.316
N	*317*	*591*

**coefficient (t-statistic)*

According to the Lee model, our stereotypical average male is estimated to earn £3.92 from the participation column compared with £4.09 from the non-participation column. The corresponding rates for a female respondent with an average set of characteristics according to the Logit/OLS estimates were £3.44 and £3.41 per hour, again a marginally positive impact. Although the derived figures are only marginally different from the Probit/OLS estimates, the difference is sufficient in the case of males, to reduce the negative impact of training participation from 5.3 to 4.8 percent.

A restriction in the models applied so far is that YT may be associated with possible interaction effects between observed and unobserved personal characteristics. One way to allow for this is to estimate a model with stronger separation; here we employ the Endogenous Switching Regression model with selection (equations 3.17a to 3.18b).[162] In this instance we assume that at any observation only one regime is observed. According to this model the estimated wage rate of a male participant with average characteristics would be £3.86 per hour compared with £4.16 per hour if they did not participate. This translates as a negative impact of 7.2 percent, larger but consistent with the direction of the other two-equation selection models. In contrast the estimated impact of participation on female youth derived from the switching model works in the opposite direction to the Probit and Logit/OLS selection models. The corresponding wage rates for a female respondent with an average set of characteristics are estimated at £3.51 for participants and £3.56 per hour for non-participants, a negative impact of approximately 1.5 percent.

---

[162] Among the canned procedures in LIMDEP, the mover/stayer model needs only minor adjustments in order to conform to the endogenous switching regression with selection, see Greene (1995, pp. 672-678).

Table 6.9: Endogenous Switching Regression wage equations for school leavers in employment, Autumn 1993

Descriptive variables	*YT*		*Non-YT*
Personal characteristics			
Female	-0.094 (2.34)		-0.157 (4.77)
Marital status	0.094 (1.39)		-0.001 (0.02)
School experience			
Unqualified	-0.047 (0.62)		-0.052 (0.55)
1-3 O/S-Grades	0.013 (0.25)		0.069 (1.14)
4+ O/S-Grades	0.047 (0.78)		0.074 (1.27)
Highers	0.107 (1.31)		0.123 (2.02)
S4 Leaver	0.166 (1.11)		0.125 (2.39)
S5 Leaver	0.126 (0.88)		0.072 (1.64)
Labour market experience			
Local unemployment rate	-0.002 (0.40)		-0.010 (1.98)
Unemployment duration (mths)	-0.007 (1.71)		-0.008 (2.90)
Hours worked	-0.012 (4.80)		-0.010 (6.76)
Does Training	0.077 (1.91)		0.026 (0.72)
Self Employed	-0.021 (0.10)		0.034 (3.94)
Industrial and occupational descriptors			
Mining and utilities	-0.039 (0.00)		0.224 (1.76)
Manufacturing	0.089 (0.74)		0.070 (0.72)
Construction	0.033 (0.27)		-0.107 (1.05)
Services	-0.058 (0.51)		-0.039 (0.42)
No industry given	-0.542 (0.00)		-0.101 (0.55)
Unskilled	-0.059 (0.97)		-0.178 (4.67)
Skilled Manual	-0.108 (2.11)		-0.079 (1.85)
Professional/Intermediate occupation	0.163 (2.21)		-0.030 (0.50)
No occupation given	0.434 (3.04)		0.056 (0.50)
Youth Training			
YT One (< 12 months)	0.096 (1.98)		..
YT Three (> 24 months)	-0.024 (0.52)		..
YT Employer	-0.082 (2.01)		..
YT Qualification	0.073 (1.79)		..
Still on YT 1993	-0.037 (0.77)		..
Constant	6.355 (25.42)		6.455 (42.28)
Log-likelihood		-745.89	
Variance parameters			
Sigma (0)		0.321 (20.25)	
Rho (0, *u*)		0.234 (0.710)	
Sigma (1)		0.285 (11.32)	
Rho (1, *u*)		-0.325 (0.94)	
N		*908*	

**coefficient (t-statistic)*

Table 6.10 summarises the estimated earnings impact across the different models in terms of the wage rate for an individual with average characteristics by gender and different aspects of YT experience. The estimated impact can be derived for the different YT options by comparing the wage rate in the relevant row of the table with the corresponding rate in the non-participant row. Focussing first on the predicted wage rates for our stereotypical mean individual (**in bold**) the pattern across the different models is fairly mixed. For male youth, the estimated impact of YT according to the base equation, which was derived using conventional OLS techniques in section 6.3.3, is minus 3.7 percent. A negative impact is also observed in the case of the two equation selectivity models ranging from minus 4.8, in the case of the Logit/OLS, to minus 7.2 percent, in the case of the switching regression, all notably larger than uncorrected OLS estimate. In contrast, the single equation treatment effects model resulted in a positive impact of approximately 5.8 percent.

In the case of female youth, a negative impact (approximately minus 3.7 percent) is again observed according to the estimates derived from the OLS equation but the direction of impact is reversed in all of the selectivity models except for the switching regression. The predicted impact of YT on the wages of a young woman conforming to an average set of characteristics ranges from marginally positive (less than one percent) in the Probit/OLS and Logit/OLS models to just over 6 percent in the Treatment Effects model. In contrast, the predicted impact derived from the switching regression is approximately minus 1.5 percent.

The inconsistencies in the derived impact estimates prevent the drawing of any general conclusions on the impact of YT on an individual with an average set of characteristics drawn from the fourth cohort of the SYPS. They do, however, illustrate that the estimates using UK data are sensitive to model choice and that this is a likely explanation for much of the variation observed in earlier studies.

Table 6.10: Estimated earnings summarised by model and YT status

*(£/hour)**	*OLS*	*Probit/OLS*	*Treatment Effects*	*Logit/OLS*	*Switching Regression*
Males					
Average YT particpant[1,2]	**3.93 (3.09-4.99)**	**3.77 (3.12-4.55)**	**4.15 (2.76-4.75)**	**3.78 (3.13-4.57)**	**3.86 (3.19-4.67)**
YT training only	3.61 (2.73-4.78)	3.50 (2.80-4.38)	3.82 (2.43-4.54)	3.51 (2.81-4.40)	3.59 (2.86-4.49)
Short spell	4.40 (3.68-5.26)	4.15 (3.60-4.79)	4.65 (3.99-5.41)	4.16 (3.57-4.86)	4.25 (3.68-4.90)
Long spell	3.78 (2.93-4.89)	3.68 (3..01-4.49)	4.00 (3.20-5.00)	3.69 (3.02-4.51)	3.77 (3.08-4.60)
YT Employer	3.58 (2.70-4.76)	3.47 (2.77-4.36)	3.79 (2.95-4.86)	3.48 (2.77-4.38)	3.55 (2.83-4.47)
Average Non-particpant[2]	**4.08 (3.27-5.09)**	**3.98 (3.21-4.94)**	**3.91 (3.09-4.94)**	**3.97 (3.20-4.93)**	**4.16 (3.41-5.09)**
Females					
Average YT particpant[1,2]	**3.41 (2.69-4.34)**	**3.43 (2.72-4.33)**	**3.62 (2.76-4.75)**	**3.44 (2.85-4.16)**	**3.51 (2.77-4.43)**
YT training only	3.14 (2.37-4.35)	3.19 (2.44-4.17)	3.32 (2.43-4.54)	3.20 (2.56-4.00)	3.26 (2.74-4.49)
Short spell	3.83 (3.20-4.57)	3.78 (3.14-4.56)	4.04 (3.25-5.03)	3.79 (3.25-4.83)	3.86 (3.20-4.66)
Long spell	3.29 (2.54-4.25)	3.35 (2.63-4.28)	3.48 (2.61-4.65)	3.36 (2.75-4.11)	3.42 (2.68-4.37)
YT Employer	3.11 (2.35-4.14)	3.16 (2.41-4.15)	3.30 (2.41-4.52)	3.17 (2.53-3.99)	3.23 (2.46-4.25)
Average Non-particpant[2]	**3.54 (2.61-4.32)**	**3.42 (2.56-4.56)**	**3.40 (2.51-4.59)**	**3.41 (2.55-4.55)**	**3.56 (2.70-4.69)**

1. *The average participant spent between 12 and 24 months on the programme and obtained a qualification from that programme.*
2. *An individual with average characteristics left school on completion of S4 with 1-3 O/S levels. They reside in an area with an unemployment rate of 9 percent, have experienced 2 months unemployment and work in the service sector at a skilled-manual occupation.*

* *Figures in parentheses give 95 percent confidence interval*

Significantly, the considerable variation in both the size and direction of the estimates, observed across the different models in table 6.10, strongly suggests that sample selection bias is inherent in the SYPS data. This is consistent with the findings of non-experimental evaluations conducted on the CETA programme in the United States in the late 1980s (detailed in section 4.3.2, see also Barnow, 1987 or Dickinson, et al. 1987). Sample selection bias was not however indicated as being a significant factor, when examining the sample selection term (Lambda) in the standard selectivity model (Heckman two-step). Furthermore, the variation observed in the direction of impact by gender across the model estimates clearly illustrates the difficulty the analyst faces when attempting to choose between different selectivity models.

In contrast, when we examine the impact of following different options within the YT programme, the derived estimates exhibit a greater consistency. For example, regardless of the model employed, the best outcome is associated with a short spell on YT (less than 12 months) when the outcome is estimated in terms of post-programme wage rates.[163] The number of participants who fit this profile are however, relatively few at approximately 15 percent of all participants. On the other hand, remaining in the employment of the employer who sponsored the YT placement consistently results in the lowest estimated wage rate. A similar portion of the YT sample remained with their YT employer. For each of these profiles the pattern is maintained for both males and females. The consistency in direction of impact with respect to shorter training events associated with YT may suggest that employers are exploiting the programme to cream the best candidates to the extent that a short-spell still resulting in a qualification reflects trainability. At the other extreme the negative impact associated with remaining in the employment of the YT sponsor may suggest that the training delivered within YT was part of a longer course or lower wage expectations among participants. This agreement across the models for these YT options suggests that it is worthwhile investigating other profiles.

The focus so far has been on an average SYPS respondent as previously defined (see also notes to table 6.10) subject to various options within YT. One of the advantages of the econometric approach, however, is that it allows the impact of training on different 'types' of individual to be evaluated. This is particularly advantageous where the objectives of the intervention include the assistance of certain target groups, such as the disadvantaged or long term unemployed. The

[163] A caveat is warranted with respect to this finding, in that the strongly positive results associated with short training spells may be due to yet another form of selectivity bias due to non-random exit from the programme. That is, those individuals that leave the programme early may have alternatives that are not observed by the analyst.

advantaged/disadvantaged profile has been adopted by a number of analysts, see for example Main and Shelly (1990) or O'Higgins (1994), and a similar approach is adopted here.

An advantaged individual is defined as male, single and having left school after completing S5 with at least one Higher. They reside in an area with an unemployment rate of 4 percent (low decile) and spend 38 hours a week at work in the service sector in a skilled manual occupation. In contrast, a disadvantaged individual is defined as male, single and having left school after completing S4 with 1 to 3 O/S levels. They reside in an area with an unemployment rate of 15 percent (high decile) and also spend 38 hours a week at work in the service sector, but in an unskilled occupation. Since leaving school they have also experienced 4 months of unemployment. Average YT experience is again assumed to be from 12 to 24 months and to result in a qualification. The results of applying these profiles to the coefficients in tables 6.4 and 6.6 to 6.9 and deriving the estimated impact, in terms of hourly wage rates, are summarised in table 6.11.

Inspection of the upper portion of table 6.11 reveals that the pattern of estimates for advantaged individuals follows that of male youth with an average set of characteristics as reported in table 6.10. In contrast, for male youth fitting the disadvantaged profile, the impact estimates are all positive following correction for the participation decision. In the case of males having had the average experience of YT, that is remaining on the programme from 12 to 24 months and procuring a qualification, the estimated impact (**in bold**) ranges from 6.3 to 17.2 percent compared with minus 3.7 percent when the 'uncorrected' OLS equation is used. Moreover, the selectivity models generally indicate that non-participation would be the least favourable option for individuals fitting the disadvantaged profile. While this is potentially a highly favourable finding from a policy perspective the relatively wide confidence intervals reported in table 6.11 indicate that the point estimates are not statistically robust and hence, at best, are only indicative.

Table 6.11: Estimated earnings summarised by model and advantage/disadvantage profiles (male youth)

*(£/hour)**	*OLS*	*Probit/OLS*	*Treatment Effects*	*Logit/OLS*	*Switching Regression*
Advantaged[1]					
Average YT particpation[2]	**4.10**	**4.09**	**4.48**	**4.07**	**4.19**
	(3.23-5.21)	**(3.39-4.95)**	**(3.41-5.87)**	**(3.37-4.92)**	**(3.47-5.06)**
YT training only	3.77	3.76	4.11	3.78	3.89
	(2.85-4.98)	(3.01-4.70)	(3.01-5.62)	(3.02-4.72)	(3.11-4.87)
Short spell	4.60	4.51	4.98	4.48	4.61
	(3.84-5.50)	(3.91-5.20)	(4.26-5.79)	(3.88-5.17)	(3.99-5.33)
Long spell	3.95	3.99	4.31	3.97	4.09
	(3.05-5.11)	(3.27-4.87)	(3.44-5.39)	(3.25-4.85)	(3.35-5.00)
YT Employer	3.74	3.77	4.08	3.75	3.86
	(2.82-4.96)	(3.00-4.74)	(3.18-5.24)	(2.98-4.70)	(3.07-4.85)
Average Non-participant	**4.26**	**4.38**	**4.21**	**4.37**	**4.46**
	(3.42-5.32)	**(3.53-5.44)**	**(3.33-5.32)**	**(3.22-5.90)**	**(3.65-5.46)**
Disadvantaged[1]					
Average YT particpation[2]	**3.60**	**3.67**	**3.85**	**3.89**	**3.93**
	(2.83-4.57)	**(3.03-4.43)**	**(2.94-5.04)**	**(3.22-4.71)**	**(3.25-4.75)**
YT training only	3.31	3.41	3.54	3.62	3.65
	(2.50-4.38)	2.73-4.26)	(2.59-4.83)	(2.89-4.53)	(2.92-4.57)
Short spell	4.03	4.04	4.26	4.29	4.33
	(3.37-4.82)	3.50-4.66)	(3.66-4.97)	(3.55-4.96)	(3.74-5.00)
Long spell	3.46	3.58	3.70	3.80	3.84
	(2.68-4.48)	(2.92-4.38)	(2.96-4.64)	(3.11-4.65)	(3.13-4.70)
YT Employer	3.31	3.38	3.51	3.59	3.62
	(2.49-4.39)	(2.69-4.08)	(2.73-4.50)	(2.86-4.51)	(2.88-4.55)
Average Non-particpant[2]	**3.74**	**3.28**	**3.62**	**3.32**	**3.48**
	(3.00-4.66)	**(2.65-4.08)**	**(2.86-4.58)**	**(2.44-4.52)**	**(2.85-4.26)**

1. An individual with advantaged characteristics is male, single and left school on completion of S5 with at least one Higher. They reside in an area with an unemployment rate of 4 percent, and spend 38 hours per week at work in the service sector in a skilled-manual occupation. An individual with disadvantaged characteristics is male, single and left school on completion of S4 with 1-3 O/S levels. They reside in an area with an unemployment rate of 15 percent, have experienced 4 months unemployment and spend 38 hours per week at work in the service sector in an unskilled occupation.

2. The average participant spent between 12 and 24 months on the programme and obtained a qualification from that programme.

** Figures in parentheses give 95 percent confidence interval*

Of course the analyst may also influence other aspects of the evaluation process that may consequently affect the outcome estimates. One area previously mentioned is that of sample construction and the common practice of constraining the sample to only include those respondents who conform to a predetermined profile. For example a number of studies focus solely on those who leave school at the earliest possible opportunity, see for example Main (1991). Such constraints are generally justified on grounds of alignment with policy, such as when the programme gives priority to 16 year olds, and/or the validity of the control group (non-participants). This latter point suggests that those who remain at school in order to attempt further examinations (Highers or A-levels) are following a different route in the school-to-work transition.[164] Given that a larger proportion of students are now staying-on at school, for at least one more year, the impact of adopting this approach is a considerable reduction in sample size. The foregoing analysis consequently included various descriptors to control for staying-on at school and the gaining of higher qualifications. However, in order to obtain an approximate idea of the impact due to the imposition of such data constraints, the OLS and selectivity equations were re-estimated with two sub-samples. The first is constrained to include only those respondents that left school before Christmas 1990, earlier defined as S4 leavers, includes just 398 cases. The second additionally includes those that left school up to a year later, or S5 leavers, increasing the sample size to 741 cases.

The restricted datasets naturally impact on the validity of some of our variables and consequently, in the Christmas 1990 leaver case, those variables indicating the attainment of one or more Highers and the relevant school leaving year are excluded from the outcome equation. While in the Christmas 1991 leaver case, the variable denominating S5 leaver is naturally excluded. The remaining variables are the same as in the full sample case. The full results for each of the models can be found in the appendix, tables A6-A10, but the results of this exercise are summarised below in table 6.13. The wage rates reported in the table again relate to male youth with an average set of characteristics as previously defined (**in bold**) together with the corresponding estimates due to different experiences of YT.

[164] Highers and A-levels are the primary entry requirement for university, respectively in Scotland and the remainder of the UK, hence 'staying on' is often called the 'academic route'; see for example Payne (1994).

Table 6.12: Estimated earnings summarised by model and different school-leaver sub-samples (male youth)

*(£/hour)**	*OLS*	*Probit/OLS*	*Treatment Effects*	*Logit/OLS*	*Switching Regression*
Christmas 1990 (S4 Leavers)					
Average YT particpation[2]	**3.84**	**4.21**	**3.60**	**4.28**	**4.39**
	(2.83-5.22)	**(3.24-5.47)**	**(2.47-5.26)**	**(3.33-5.51)**	**(3.40-5.68)**
YT training only	3.50	3.90	3.29	3.78	4.09
	(2.71-5.42)	(2.89-5.25)	(2.18-5.00)	(2.97-5.29)	(3.06-5.47)
Short spell	4.33	4.60	4.06	4.68	4.82
	(3.38-5.56)	(3.70-5.73)	(2.95-5.59)	(3.79-5.78)	(3.90-5.96)
Long spell	3.54	3.86	3.32	3.93	4.07
	(2.51-5.00)	(2.86-5.23)	(2.20-5.03)	(2.93-5.27)	(3.04-5.46)
YT Employer	3.70	4.01	3.47	4.08	4.18
	(2.67-5.11)	(302-5.33)	(2.33-5.15)	(3.10-5.37)	(3.15-5.53)
Average Non-participant	**3.98**	**4.30**	**4.28**	**4.33**	**4.94**
	(2.96-5.33)	**(3.14-5.90)**	**(3.18-5.78)**	**(3.16-5.92)**	**(3.52-6.74)**
Christmas 1991 (S4 & S5 Leavers)					
Average YT particpation[2]	**3.98**	**3.74**	**4.35**	**3.75**	**3.87**
	(3.08-5.15)	**(2.95-4.75)**	**(3.52-5.34)**	**(2.96-4.76)**	**(3.06-4.91)**
YT training only	3.68	3.48	4.01	3.49	3.60
	(2.74-4.94)	(2.65-4.56)	(3.19-5.06)	(2.66-4.58)	(2.75-4.72)
Short spell	4.47	4.11	4.86	4.12	4.24
	(3.65-5.47)	(3.39-4.98)	(4.13-5.71)	(3.40-5.00)	(3.50-5.15)
Long spell	3.81	3.64	4.17	3.65	3.77
	(2.88-5.03)	(2.84-4.67)	(3.30-5.25)	(2.84-4.70)	(2.94-4.84)
YT Employer	3.74	3.48	4.08	3.49	3.60
	(2.81-4.98)	(2.65-4.56)	(3.21-5.20)	(2.67-4.56)	(2.74-4.72)
Average Non-particpant[2]	**4.14**	**4.10**	**3.87**	**4.11**	**4.11**
	(3.26-5.26)	**(3.29-5.10)**	**(3.01-4.97)**	**(3.12-5.37)**	**(3.08-5.38)**

1. The average participant spent between 12 and 24 months on the programme and obtained a qualification from that programme.
2. An individual with average characteristics left school on completion of S4 with 1-3 O/S levels. They reside in an area with an unemployment rate of 9.5 percent, have experienced 2 months unemployment and work in the service sector at a skilled-manual occupation.
** Figures in parentheses give 95 percent confidence interval*

With respect to Christmas 1990 leavers, in the case of the average male youth, the models consistently associate a negative impact with participation in YT. Considerable variance is observed in the magnitude of the derived impact estimates, however, which range from marginally negative in the case of the Logit/OLS to as much as minus 15.8 percent in the case of the treatment effects model. In general, this is again offset if the participant only remains on YT for a short period (less than 12 months) yet still gains a qualification. The selection models indicate that remaining with a YT employer or prolonging duration on YT (greater than 24 months) are the least favourable options for early school leavers.

The relatively higher wage rates, observed for the early school leavers in the upper portion of table 6.13, are consistent with the empirically and statistically significant coefficients of the variable 'S4 leaver', which indicated early leaving in the full sample equations. This group naturally has the highest level of labour market experience and this could be a reflection of this factor. It should be noted, however, that this sample includes the highest hourly wage rate reported in the SYPS earnings sample (£16.88 compared to the sub-sample average of £4.02) and the influence of this might be disproportionate in such a small sample. However, when the equations were re-estimated excluding several of the more extreme values there was no apparent difference in the derived impact.

When Christmas 1991 school leavers were added to the sample the resultant pattern in the impact estimates conforms to that observed in the full sample case. The single equation treatment effects model is again at variance with the other models in predicting a positive impact of 11 percent. In contrast, the remaining selectivity models predict an impact of between minus 8.7 percent and minus 10 percent. These are all significantly larger than the estimates derived from the full sample (respectively plus 5.8 and minus 4.8 to 7.2 percent). On the other hand, the OLS estimate of minus 3.8 percent is almost identical to the full sample estimate of 3.7 percent.

Clearly, a wide range of restrictions can be placed on the data in order to either focus on specific types of participant, aspects of the programme and/or in order to construct a more acceptable control group. For example YT participation might be constrained to include only those that entered it immediately after leaving school or the comparison group to those who experience unemployment but opt not to partake in YT (eligible non-participants). It is important to note, however, that in both of the sub-samples analysed here, that the reported confidence intervals are considerably wider than those observed for an average individual in the full sample case. This indicates that, as might be expected, the reduction in sample size is accompanied by an additional trade-off in the form of a drop in statistical significance and the subsequent results are again rendered, at best, indicative. Moreover, in spite of the parameters placed on the dataset, the basic dichotomy of all YT participants compared to all non-participants is maintained. An alternative means of focussing on a specific group and allowing comparison with other like groups is to adopt the finer level of typology available in a polychotomous choice model (equations 3.25 to 3.31 in section 3.5.2). Within this class of models, the Multinomial Logit is the most widely employed to explore the impact of various training options within the school-to-work transition, see for example Dolton et al., (1994b, 1995); Green et al., (1996); and Payne (1995). The relative brevity of coverage in this cohort of the SYPS invalidates a number of the options used in these existing studies, such as apprenticeships, as the respondents would still be immersed in such programmes. Nevertheless, it is possible to examine a more limited selection of options in order to assess the sensitivity of impact estimates to the addition of these choices in the school-to-work transition.

The method adopted here is to estimate the gross hourly earnings equations for employees who followed one of five possible routes into the labour market. The sample remains as previously described including all the 908 cases of the earnings sample. The five options distinguished are:

(0) Those respondents who entered the labour market at the earliest possible opportunity (aged 16) and did not take any further education of training within the survey period. This is our reference category, typically called the 'Nothing' route;

(1) Those who received some post-compulsory education but no further training defined as the 'Further Education' route;

(2) Those who entered into YT scheme which was not designed to lead to qualifications (or they were not awarded any qualification);

(3) Those who entered YT designed to lead to a qualification, which was successfully obtained;

(4) Those respondents who indicated that they received company training that involved a course or led to a qualification.

An attempt has been made to capture within each group a relatively homogenous education and training experience but clearly, these trajectories condense a large number of possible routes in the school-to-work transition. Choice is limited by the available data, the need to maintain degrees of freedom within each group and the length of the survey. Nevertheless, the advantage of this approach is readily apparent even within YT given it allows us to make a distinction between those course that lead to a qualification and those that do not. Participation in one or other of the trajectories is modelled in a similar fashion to before and we employ the same variables reflecting educational performance at 16 years of age and family background.[165] The results of the Multinomial Logit selection model are presented in table 6.13a.

Table 6.13a: Multinomial selection equations [1, 2]

Descriptive variables	*Further Education (no training)*	*YT (no qualification)*	*YT (qualification)*	*Company Training*
Female	0.329 (1.75)	-0.022 (0.05)	-0.53 (1.41)	-0.544 (2.48)
School Experience				
Unqualified	-2.915 (8.15)	1.933 (5.41)	0.885 (3.11)	0.862 (1.98)
1-3 O/S-levels	1.848 (7.55)	-0.925 (3.81)	0.549 (2.22)	1.165 (2.86)
4+ O/S-levels	1.664 (6.33)	-0.151 (0.46)	-0.960 (2.54)	0.709 (2.90)
Truant (habitual)	-1.395 (3.52)	0.251 (0.81)	-0.332 (1.15)	-0.700 (1.59)
Fathers occupation				
Skilled/ manual	0.315 (1.13)	-0.548 (2.09)	-0.351 (1.34)	0.172 (0.57)
Professional/intermediate	0.568 (1.83)	-1.701 (4.80)	-1.604 (1.96)	0.081 (0.26)
No occupation given	0.332 (1.03)	-0.410 (1.29)	-0.349 (1.29)	-0.426 (1.06)
Father unemployed	-0.680 (1.61)	-0.105 (0.29)	-0.538 (1.59)	-0.843 (1.66)
Local unemployment rate	0.022 (0.86)	0.084 (2.96)	0.079 (3.53)	0.008 (0.31)
Constant	-1.388 (3.70)	-1.366 (3.58)	-0.858 (6.76)	-0.758 (2.02)
N	*436*	*165*	*322*	*171*
Log-likelihood		-1801.75		
Chi-squared		355.79		

Notes:
[1]. Baseline = Nothing (199 cases)
[2]. coefficient (t-statistic)

[165] The variable indicating the type of school attended in the fourth form was however dropped due to insufficient variation across the different routes.

The results of the selection model largely conform to expectations and observations in the binomial analysis. The key influences appear to be the level of educational attainment and the occupation of the respondent's father. Higher educational attainment and having a professionally qualified father is associated with a greater likelihood of staying on at school (further education) or undertaking company training. Otherwise the respondents are more likely to participate in the YT scheme. The results also suggest that higher local unemployment rates encourage young people away from the 'do nothing' route into either of the YT options. The model also indicates that female youth are less likely to be involved in company training and more likely to stay on at school, otherwise there seems to be a relatively weak connection between the remaining personal characteristics and participation in YT.

The limitation of the relatively small SYPS earnings data sub-sample for multinomial analysis can be seen in table 6.13b, which reports the requisite earnings equations. To maintain degrees of freedom and avert collinearity problems the number of descriptive variables is reduced to just those detailing personal characteristics, schooling experience (with the exception of highers and school leaving year), labour market experience (without the *does training* variable), and the industrial/occupational descriptors. This restricts the construction of individual profiles that match those used to generate the hourly earnings estimates derived from the binomial models above. Even in this restricted form, few of the variables were subsequently estimated with significant coefficients and some of the signs were also counter-intuitive, suggesting that the results should be treated, at best, as only indicative. The predicted earnings of an individual with approximately average characteristics range from £3.21 per hour, if the individual does not participate in either further education or training, up to £4.18 per hour, if they follow the company-training route. YT participation results in an estimated rate of £3.27 per hour if qualifications are part of the scheme, but a comparatively poor rate of £3.03 per hour if qualifications were not attempted. Finally, following the further education route resulted in predicted earnings of £3.21 per hour.

Table 6.13b: Multinomial Logit earnings equations [1]

Descriptive variables	*Do Nothing*	*Further Education*	*YT (no qual.)*	*YT (qualification)*	*Company Training*
Personal					
Female	1.680 (1.18)	0.508 (1.18)	-0.176 (0.26)	0.077 (0.11)	1.230 (0.96)
Marital Status	0.120 (1.42)	-0.325 (1.42)	0.128 (1.02)	0.934 (0.60)	0.099 (0.36)
School Experience					
Unqualified	-0.132 (0.88)	0.194 (0.83)	0.072 (0.35)	-0.178 (0.13)	-0.788 (0.27)
1-3 O/S-levels	-0.336 (0.23)	0.120 (0.73)	-0.184 (0.87)	-0.254 (2.14)	0.136 (1.42)
4+ O/S-levels	0.223 (0.14)	-0.328 (0.19)	0.261 (0.82)	0.614 (0.44)	0.223 (0.98)
Labour market					
Local unemployment	-0.031 (1.13)	-0.006 (0.68)	0.025 (0.68)	-0.018 (0.93)	-0.026 (0.18)
Unemployment (mths)	0.568 (1.83)	0.007 (0.92)	-0.014 (1.04)	-0.254 (2.52)	-0.153 (0.16)
Hours worked	-0.032 (2.03)	-0.017 (2.89)	0.008 (2.81)	0.089 (0.58)	0.009 (4.13)
Self employed	-0.680 (1.61)	-0.562 (1.37)	-0.279 (1.50)	-0.201 (0.51)	0.318 (0.64)
Industry/occupation					
Mining	-0.141 (0.32)	0.336 (0.85)	0.001 (0.00)	-0.319 (0.39)	-0.265 (0.59)
Manufacturing	0.484 (0.17)	0.298 (1.08)	0.167 (0.05)	0.120 (0.38)	0.337 (1.84)
Construction	0.865 (0.29)	0.285 (0.98)	-0.155 (0.15)	0.129 (0.40)	0.279 (0.68)
Services	-0.107 (0.40)	0.628 (2.50)	-0.097 (1.50)	0.138 (0.46)	0.132 (1.35)
No industry given	0.858 (0.34)	0.067 (0.23)	-0.153 (0.80)	-0.197 (2.08)	0.105 (0.01)
Unskilled	1.49 (0.92)	0.067 (1.79)	-0.191 (0.51)	0.089 (0.58)	0.526 (0.27)
Skilled Manual	-0.161 (0.09)	-0. 201 (1.17)	0. 021 (0.05)	-0. 113 (0.81)	-0. 128 (1.78)
Professional/int.	-0.016 (0.34)	-0.279 (1.65)	-0.133 (2.27)	0.119 (0.09)	0.185 (1.09)
No occupation	0.132 (0.46)	-0.230 (1.35)	-0.377 (0.61)	-0.511 (0.13)	-0.418 (1.35)
Lamda	0.401 (1.06)	0.347 (1.46)	0.347 (1.46)	0.547 (1.77)	0.210 (0.34)
Constant	5.811 (1.13)	5.659 (3.70)	5.466 (0.26)	3.710 (0.93)	6.076 (6.49)
R^2	0.141	0.232	0.133	0.288	0.317
N	*199*	*436*	*165*	*322*	*171*

Notes:
[1]. coefficient (t-statistic)

The application of the different selectivity models to the SYPS data has clearly illustrated that earnings impact estimates are highly sensitive to both model choice and sample construction. The selectivity models do appear to be good indicators of the direction of impact for various options and profiles but for the majority of participants the early returns to training would appear to be negative. Notable exceptions are the positive impact associated with YT events that result in a qualification but last less than 12 months and/or for individuals fitting the disadvantaged stereotype, but these benefited only a small portion of the sample and rested on statistically relatively weak estimates.

6.4: Estimated Impact of Youth Training on Employment

The other key economic outcome used in the evaluation literature is post-programme employment. At first glance this measure would appear to fit more comfortably with the objectives of the YT intervention, which is primarily designed to facilitate the school-to-work transition (see discussion in chapters 2 and 4). The programme has largely evolved out of the earlier manpower programmes beginning with the YOP, but the emphasis on training has become more pronounced with the development and support of subsequent intervention. As such, YT can be viewed as an alternative route to the traditional direct entry route of employment. Moreover, the point in time analysis typically adopted in evaluations does not reveal anything about the quality or longevity of any jobs gained. Again post-programme employment, although clearly a desirable outcome, can only be considered an approximate measure of programme effectiveness. In general, existing research indicates that participation in government-assisted youth training programmes is associated with a positive impact on the subsequent employment prospects of participants.

This section reports the results of analysis on the estimated employment probability of a sample of respondents drawn from the fourth cohort of the SYPS beginning with the Probit model. This is preceded by a preliminary examination of the employment data in the SYPS and a discussion on the factors involved in the determination of employment.

6.4.1: Employment Data in the SYPS

The dependent variable is employment status (employed or not) in October 1993 when the follow-up sweep of the SYPS was conducted. Employment status was determined from the survey questions on the respondents' 'main activity' at that time. The survey question pertaining to the individual's main activity was structured in the same format as the diary and simply asked, "What are you doing now?" The options were: on youth training, in a full-time job *but not on YT*, unemployed and looking for work, in full-time education at school, in full-time education at college or university, or alternatively 'doing something else'. The last option asked them to clarify what they were doing. Responses were then cross-checked with other survey questions which specifically asked the respondent to indicate if they were unemployed, in paid employment (including part-time) and other aspects of work. The resultant employment sample comprises 1,293 cases representing all the SYPS respondents who were identified as being in the labour market of which 1,044 (approximately 80 percent) were in employment (see also figure 6.1).

In table 6.14, descriptive statistics on a selection of variables are presented for all sample respondents in the labour market in October 1993. The table provides the means and (where relevant) standard deviations of the variables included in the analysis. The first two columns of the table report the statistics separately for those respondents who are in employment and those who are unemployed while the third column presents the statistics for the sample as a whole. This allows a rough indication of the likely effect of the different variables in determining the employment status of the young person at the time of the follow-up sweep. To aid clarity the explanatory variables are again grouped into five broadly related categories. Many of the variables are common to the earnings analysis but here the focus is on variables that might influence employment in terms of opportunity, choice and attractiveness to potential employers.

In the first category, *personal characteristics*, the controls for gender and marital status are retained. The impact of gender on employment probability has varied across studies examining the earnings impact of youth training programmes in the UK. For example, being female was associated with higher employment probability in the study conducted by Green, et al (1995), while the studies of Main and Shelly (1990) and O'Higgins (1994) both identified a negative effect. The descriptive statistics in table 6.14 indicate that a positive effect is likely in this case because a relatively greater proportion of young women are included in the employment column (column one) than in the unemployed column (column two). Being married is again expected enhance the probability of being in employment due to the perception of stability and willingness to undertake responsibility although there is little difference apparent in the relevant descriptive statistics.[166]

In the *school experience* category, the same measures of educational attainment are carried over from the previous section, with the omitted class in the analysis again being 'O/S-levels with D & E grades only'. A cursory inspection of table 6.14 suggests that better educational performance is associated with a higher probability of being in employment. This pattern is reinforced by the TOTSCEP variable (an average of 6.07 for those in employment compared with 4.17 for the unemployed), although it is again only listed for reference purposes and is not included in analysis below. A larger proportion of the S4 school-leavers are in unemployment while the opposite case is observed for S5 leavers. Two additional variables are introduced into the analysis in an attempt to indicate aspects of the respondents' initiative and motivation. A paid part-time job while at school is likely to have a positive impact on the chances of finding work after leaving school, to the extent that it reflects initiative and motivation. It also indicates early experience of the labour market although it is not possible to separate these effects. As in the participation Probit (section 6.2.2), a dummy variable reporting 'habitual' truancy at school is included and it is expected that it will be negative in its influence on the probability of gaining employment to the extent that it reflects the individual's motivation.

Variables are also included to represent *family background*. As in the modelling of the participation decision, these mostly relate to the employment and occupational status of the individual respondent's father at the time of the first sweep. The first is derived from survey questions on the father's reported employment status and the next four are from questions on their reported occupational classification, with the omitted class being 'manual'. We would expect having a father in a professional/intermediate or skilled/manual occupation to be associated with a positive influence on the employment probability of a respondent to the extent that

---

[166] I also constructed a dummy variable for individuals identified as belonging to an ethnic minority (equal to 1 if the respondent described themselves as being African, Caribbean, Indian, Pakistani, or Bangladeshi). However, the majority of respondents remaining in this category after the follow-up sweep (45 cases) continued on into higher education leaving an unrepresentative number (6) in the sub- sample.

they represent the unmeasured influences associated with a positive socio-economic background. Table 6.14 indicates that this appears to be the case with a relatively larger proportion of respondents in employment for both variables. On the other hand, having an unskilled or unemployed father is expected to be negatively related to the respondents' likelihood of finding work, reflecting constraints on the fathers' labour market knowledge which is likely to be transmitted to their children. The increased labour market activity rates for women in recent decades (as observed in chapter two) suggest that a mother's employment and occupational status may be an equally important influence on a young person's employment prospects. A variable indicating if the mother was in a non-manual job is hence included and inspection of the table reveals that it appears to be associated with a significant positive impact on respondents' probability of employment. A further dummy variable is included to reflect if the respondent was living in an owner occupied household at the time of the first sweep as an additional means of taking into account the social status of the individual.

The local unemployment rate and unemployment duration are again included in the *labour market experience* category with the addition of two variables indicating the individual's labour market status at two intermediate dates, specifically October 1991 and October 1992. Early experience of unemployment may have a mixed impact on subsequent employment probability. On the one hand, the experience could motivate the individual to widen their search or improve their suitability for employment or, on the other hand, it may have a scarring effect lowering their confidence and/or reducing their attractiveness to a potential employer. Employment status soon after leaving school is also an important explanatory variable, in as much as YT was designed as an alternative route in the school-to-work transition and not just to absorb excess youth unemployment. More recent experience of unemployment is expected to have a negative impact on current prospects.

The final category of variables includes those pertaining to the respondents' *youth training experience* between leaving school and October 1993 and includes the same variables as in the earnings case with the exception of 'YT employer', which naturally becomes redundant.[167] An additional control is, however, introduced to indicate the number of youth training events/courses that the individual participated in. For this category the means of the descriptors listed in table 6.14 (*in Italics*) only include those who participated in YT (487 cases). This gives an initial rough indication of the apparent impact of YT participation on subsequent employment.

[167] The omitted class among the training duration variables is again 12 to 24 months.

Table 6.14: Descriptive statistics on independent variables, employment*

In the labour market Autumn 1993	Employed *(N = 1,045)*	Unemployed *(N = 248)*	Whole sample *(N = 1,293)*
Personal characteristics			
Female	0.55	0.48	0.54
Married	0.07	0.08	0.07
School experience			
Unqualified	0.10	0.27	0.13
O/S-levels with D & E grades only	0.14	0.22	0.15
One to three O-levels	0.24	0.24	0.24
Four or more O-levels	0.19	0.09	0.17
One or more Highers	0.33	0.18	0.30
TOTSCEP	6.07 (3.72)	4.17 (3.57)	5.71 (3.76)
Part-time job	0.62	0.41	0.58
Truant (habitual)	0.07	0.18	0.09
S4 leaver	0.44	0.55	0.46
S5 leaver	0.38	0.30	0.36
Background characteristics			
Father unemployed	0.05	0.12	0.06
Father unskilled	0.03	0.07	0.04
Father manual	0.11	0.19	0.13
Father skilled/manual	0.45	0.35	0.44
Father professional/intermediate	0.25	0.15	0.23
Father's occupation not given	0.15	0.24	0.16
Mother in non-manual job	0.43	0.28	0.40
Owner occupier	0.60	0.43	0.57
Labour Market experience			
Local Unemployment Rate	9.02 (3.60)	10.13 (3.40)	9.23 (3.59)
Duration of Unemployment (mths)	4.21 (3.97)	10.41 (9.41)	3.61 (6.36)
Unemployed October 1991	0.02	0.10	0.04
Unemployed October 1992	0.04	0.32	0.09
Youth Training[1]			
YT experience	0.34	0.52	0.38
YT One (< 12 months)	*0.35*	*0.45*	*0.38*
YT Two ($\geq 12 \leq 24$months)	*0.31*	*0.30*	*0.31*
YT Three (> 24 months)	*0.33*	*0.26*	*0.31*
YT number of courses	*1.20 (0.52)*	*1.51 (0.76)*	*1.28 (0.61)*
YT qualification	*0.71*	*0.54*	*0.66*
YT 1993	*0.13*	*0.15*	*0.13*
N (YT sample)	*357*	*130*	*487*

* *Standard deviations of continuous variables in parentheses*
[1.] *YT descriptors only include those who participated in YT between leaving school and October 1993*

Participation in YT *per se* appears to have a negative effect on respondents' employment prospects. In terms of the descriptive variables detailing YT experience, the negative effect appears to be primarily associated with recent YT experience and the number of training courses the individual has been on. Significantly, short spells on YT, which resulted in the largest measure in terms of earnings impact, appear to be associated with a negative effect on post-programme employment probabilities. On the other hand, gaining a qualification as a result of the YT involvement clearly continues to have a positive effect on labour market outcomes.

It is also possible to gain an impression of the employment impact associated with participation in YT while still using the aggregate survey data by drawing on the SYPS diary information. Indeed, such information allows us to gauge apparent employment effects of YT for different groups of school-leavers over time. For example, table 6.15a illustrates the transition of sample labour market states between two survey dates for those that left school at the earliest possible opportunity, that is, at 16 years of age. This sub-sample is labelled 'S4 leavers' and includes all those respondents that left school at the end of S4 (summer 1990) or at the end of the first term of S5 (Christmas 1990). The first SYPS diary entry on subsequent labour market status is then April 1991, which is thus used as the reference point. In other words, the table (or transition matrix) shows labour market status in October 1992 as a function of labour market status in April 1991. Inspection of the transition matrix reveals that, approximately 18 months after leaving school, there is little evidence of any employment impact associated with YT participation when compared with post-school unemployment. It is also apparent, however, that a significant proportion of participants remained on the programme at this time.[168]

---

[168] The sub-samples in the transition tables differ slightly from the corresponding size of S4 and S5 leaver samples in the employment sample, as detailed in table 6.14, because they include some respondents who remain outwith the labour market or who have changed transition states. For example, the S4 leaver sample originally included 681 sample respondents, however 64 of these entered directly into higher education, 8 subsequently returned to school. A further 5 did not enter the labour market (looking after family or disabled/unable to work), which leaves a final sample of only 604 cases. Additionally, cases were only included in the transition matrixes where the requisite diary information was complete reducing the sample to 587 cases.

Table 6.15a: Labour market status in October 1992 by labour market status in 1991, S4 leavers*

		Labour Market State in October 1992 (%)			
		Employed	*YT*	*Unemployed*	*% by state in April 1991*
Labour Market State after leaving school (%)	*Employed*	88.0	4.8	7.2	41.2
	YT	45.7	39.4	14.9	48.5
	Unemployed	44.6	8.5	46.9	10.3
	% by state in October 1992	63.4	21.9	14.7	**100.00**

** Early school leavers only, i.e. those that left by Christmas 1990 or earlier, N = 587*

Table 6.15b: Labour market status in October 1993 by labour market status in 1991, S4 leavers*

		Labour Market State in October 1993 (%)			
		Employed	*YT*	*Unemployed*	*% by state in April 1991*
Labour Market State after leaving school (%)	*Employed*	85.6	2.1	12.3	41.2
	YT	69.0	6.3	24.2	48.5
	Unemployed	51.1	8.5	40.4	10.3
	% by state in October 1993	73.2	4.8	22.0	**100.00**

** Early school leavers only, i.e. those that left by Christmas 1990 or earlier, N = 587*

Table 6.15c: Labour market status in October 1993 by labour market status in 1992, S5 leavers*

		Labour Market State in October 1993 (%)			
		Employed	*YT*	*Unemployed*	*% by state in April 1992*
Labour Market State after leaving school (%)	*Employed*	89.1	0.1	10.8	62.3
	YT	66.4	15.9	17.7	28.0
	Unemployed	51.3	5.4	43.3	9.7
	% by state in October 1993	79.3	5.0	15.7	**100.00**

** Includes only those respondents that left school between Christmas 1990 and Christmas 1991, N = 462*

Table 6.15b shows labour market status for the same group of early school-leavers 12 months later in October 1993, again as a function of labour market status in April 1991. Inspection of this transition matrix reveals that S4 school leavers had a significantly higher probability of being in employment in October 1993 if their post school labour market status is YT as opposed to unemployment. The table indicates that entering YT after leaving school renders a 17.9 percentage point advantage over being unemployed when first entering the labour market in terms of being in employment approximately 30 months later. A similar pattern is observed in table 6.15c, which focuses on those respondents that entered the labour market up to one year later, that is Christmas 1991 or S5 leavers. In this case the reference point is April 1992. Although the respondents had only been in the labour market for approximately 18 months, at the time the follow-up sweep was conducted, the transition matrix indicates that YT still renders an advantage in the order of 15 percentage points over post-school unemployment.

Of course, this approach only presents us with an overview of the apparent, or gross, employment effects of YT and makes no attempt to allow for what would have happened in the absence of the programme. In order to do this we need to consider the influence of all those factors that were described in this section as likely to have an impact on the employment prospects of both participants and non-participants. That is, in order to derive a more accurate measure of programme impact we need to take into account differences in the personal, family and labour market characteristics of the individual respondents as we did in the earnings case. The next section presents different methods for evaluating the impact of participation in YT in terms of post-programme employment.

6.4.2: Employment Estimates

The most common model used in evaluating the employment impact of government assisted youth training programmes is the univariate Probit,[169] which estimates the probability that a school leaver is in employment on a given reference date, usually the date of the survey's follow-up sweep. The full results from applying the Probit to

---

[169] The Probit model was presented formally in section 3.5.2, equations 3.21 to 3.24.

the SYPS employment sample are presented in table 6.16, separately for the sample as a whole (1293 cases) and the S4 leaver sample (596 cases). The latter is to allow closer approximation with some earlier studies and justified on grounds of the relative validity of the comparison group and closer alignment with policy.

Focusing first on the estimates for the whole sample in can be seen in the table that employment appears most directly influenced by the respondent's level of attainment at school and subsequent experience of unemployment. Employment probability is heightened with increasing levels of human capital accumulation as represented by higher school leaving qualifications.[170] Family background is also significant with those respondents who have a father in a skilled/manual or professional/intermediate occupation more likely to be in employment. On the other hand, those with an unemployed father or those who have experienced unemployment, particularly in October 1992, are less likely to be in employment. In terms of YT experience, the average participation of 12-24 months appears to have a positive effect although this does not reach conventional levels of statistical significance.[171] Surprisingly, the two key aspects contributing to a positive YT experience in the earnings case, that is a short training spell and the gaining of a qualification, both appear to have a small negative effect on employment probability. Similarly, recent YT experience and the number of training events reduce employment prospects with the latter statistically significant, which may mean that multiple YT events are working as a negative signalling mechanism to employers. The signs for the remaining explanatory variables largely conform to prior expectations.

---

[170] Again, it is equally feasible that employers are using academic credentials as a signalling device indicating that the individual possesses other hard to measure attributes rather than the implied human capital accumulation (Main, 1985, p. 434)

[171] In order to interpret the Probit coefficients in a similar fashion to OLS estimates, one needs to take the partial derivative of the variable of interest X_i, that is: $\frac{\partial \Pr(y_i = 1)}{\partial X_i} = \beta_k \phi(X_i \beta)$. In this case, when evaluated at the sample mean of the descriptive variables, the marginal change in the probability of being in employment in October 1993 that can be attributed to YT participation is approximately 4.8 percent in the full sample case and 4.4 percent in the S4 leaver sample. Given the artificial construct of using the mean of all the variables in the employment equation as the base, the stylised profiles presented below in table 6.17 are preferred.

Table 6.16: Probability that a school leaver is in employment in October 1993, using SYPS data*

Descriptive variables	*Whole Sample*	*S4 Leavers*
Personal characteristics		
Female	-0.001 (0.01)	-0.037 (0.25)
Married	0.244 (1.27)	0.493 (1.73)
School experience		
Unqualified	-0.129 (0.77)	-0.191 (0.97)
1-3 O/S-levels	0.171 (1.14)	0.082 (0.42)
4+ O/S-levels	0.338 (1.88)	0.369 (1.40)
Highers	0.432 (2.25)	..
S4 leaver	0.532 (2.73)	..
S5 leaver	0.281 (1.73)	..
Habitual truant	-0.295 (1.77)	-0.392 (2.05)
Part time job while at school	0.167 (1.67)	0.015 (0.10)
Background characteristics		
Father unemployed	-0.367 (2.00)	-0.486 (1.85)
Father in skilled/ manual occupation	0.319 (2.48)	0.495 (2.67)
Father in professional/ intermediate	0.215 (1.39)	0.519 (1.98)
Father's occupation not given	0.230 (1.52)	0.434 (2.09)
Mother in non-manual occupation	0.140 (1.32)	0.167 (0.96)
Labour market experience		
Local unemployment rate	-0.013 (0.91)	0.001 (0.07)
Unemployment experience (mths)	-0.089 (10.11)	-0.074 (6.60)
Unemployed October 1991	0.241 (1.00)	0.088 (0.28)
Unemployed October 1992	-0.606 (3.84)	-1.088 (5.64)
Youth Training		
YT Experience	0.194 (0.73)	0.182 (0.54)
YT One (< 12 months)	-0.050 (0.26)	-0.250 (1.02)
YT Three (> 24 months)	0.024 (0.12)	0.008 (0.03)
YT Number of courses	-0.302 (2.53)	-0.299 (2.12)
YT Qualification	-0.012 (0.07)	0.071 (0.35)
Still on YT 1993	-0.410 (1.84)	-0.659 (2.35)
Constant	0.749 (2.67)	1.186 (3.93)
Chi-square	386.86	248.48
Log-likelihood	-438.62	-197.07
N	*1,293*	*596*

**coefficient (t-statistic)*

The pattern is replicated when focussing only on S4 leavers although a number of additional parameter estimates reach statistical significance, notably for the variables indicating that a respondent acknowledged regularly playing truant while still at school and respondent's who had recent experience on YT. The results in both cases indicate that YT experience is much less important than many of the other factors that determine the probability of being in employment. Significantly, YT participation does little to mitigate some of the major factors that reduce employment chances, such as having experienced unemployment or having an unemployed father.

In order to interpret the impact of participation in YT on subsequent employment in a more readily applicable fashion, the parameter estimates in table 6.15 are applied to the profiles (expected experience) of different ideal types of school-leaver and evaluated using the values of the standard normal distribution. In this case a school-leaver with an average set of characteristics is defined as male, single and having left school on completion of S4 with one to three O/S levels. They reside in an area with an unemployment rate of 9 percent, have experienced 2 months unemployment (between April 1991 and October 1993) and have a father in a skilled-manual occupation. In addition to these characteristics, the average YT participant is again defined as having spent between 12 and 24 months on the programme and obtained a qualification from that involvement. Table 6.17 presents the resultant *expected probabilities* for this average respondent together with those fitting the advantaged/disadvantaged profiles as defined in the notes to the table.[172]

Table 6.17a: The expected probability of employment in October 1993 for different types of school-leaver by YT experience*

	Whole Sample		S4 Leaver	
	Non-YT	*YT*	*Non-YT*	*YT*
Average	0.76	0.81	0.66	0.75
Advantaged	0.83	0.88	0.80	0.87
Disadvantaged	0.30	0.37	0.07	0.10

* *An individual with average characteristics is male, single and left school on completion of S4 with 1-3 O/S levels. They reside in area with an unemployment rate of 9 percent, have experienced 2 months unemployment and their father is in a skilled-manual occupation. An individual with advantaged characteristics is male, single and left school on completion of S5 with at least one Higher. They reside in an area with an unemployment rate of 4 percent, and also have a father in a skilled-manual occupation. An individual with disadvantaged characteristics is male, single and left school on completion of S4 with no qualifications. They reside in an area with an unemployment rate of 15 percent, have experienced 4 months unemployment, were also unemployed in both October 1991 and October 1992 and their father is in a manual occupation. Among S4 leavers an advantaged individual is assumed to have 4 plus O/S grades, rather than highers (and to have left after completing S4), otherwise the profiles are the same. All those with YT experience are assumed to have spent between 12 and 24 months on the programme and obtained a qualification from that programme.*

[172] These differ slightly from the earnings case due to the different model constructs.

Taken over the whole SYPS employment sample, the effect of participating in YT is seen from table 6.17a to increase the probability of being in employment by approximately 5 percentage points in the case of either an average or an advantaged school leaver. In the case of a disadvantaged leaver YT participation is associated with increased employment probability of some 7-percentage points. Whereas when the expected probabilities are only taken over those who entered the labour market before Christmas 1990, YT participation is associated with an increase in employment probability in the order of 9 percentage points for the average S4 leaver or 7 percentage points for those conforming to the advantaged profile. In the case of the disadvantaged S4 leaver, however, YT confers an advantage of just 3 percentage points. As highlighted in the earlier studies of Main and Shelly (1990) and O'Higgins (1994) it is evident that the effect of YT is quite small in comparison with the personal circumstances and characteristics that differentiate the disadvantaged from the average and advantaged stereotypes. It would appear that an advantaged school leaver will have little trouble procuring employment whether they remain at school after the fourth form or not. On the other hand, the prospects of disadvantaged school leavers appear to be particularly poor if they leave school after the fourth form and YT can do little to mitigate this disadvantage.

As observed in section 6.3.3, where a Probit was used to model the participation decision, it is advisable to test heteroskedasticity when applying this model. Indeed, if the error term in the Probit is not homoskedastic, estimation produces not only inconsistent estimates of standard errors as in the linear model but also inconsistent parameter estimates (O'Higgins, 1994, p. 611).[173] The same procedure is used as in section 6.3.3 using an auxiliary regression including those variables, which may cause heteroskedasticity. In the first instance a general heteroskedastic model was used incorporating most of the descriptive variables in the employment equation. The derived likelihood ratio test statistic, for the full sample, was 21.29 compared with the five-percent critical $\chi^2(16)$ value of 26.29 and hence we cannot reject the

[173] This is because the error variance, σ, is not identified meaning that in practice one estimates β/σ rather than β. The normalisation, $\sigma = 1$ causes no problems as long as σ is constant, however if it is not, the variation in σ will effect the parameter estimate β (O'Higgins, op cit., p. 611-612). See also Davidson and MacKinnon (1984).

null of homoskedasticity. In the second instance the test is refined to include those labour market experience variables indicating if the individual was unemployed in October 1991 and/or October 1992, if they had YT experience and if they were still on YT in April 1993. In this case, the derived likelihood ratio test statistic of 12.78 would be narrowly accepted at the one-percent critical $\chi^2(4)$ value of 11.07 but would have led to the rejection of the null at the five-percent critical value of 9.49. As such, the expected probabilities for our different types of school leaver were recalculated using the parameter estimates corrected for heteroskedasticity, which are presented in table A11 in the appendix to this thesis. The results of this exercise are shown in table 6.17b.[174]

Table 6.17b: The expected probability of employment in October 1993 for different types of school-leaver by YT experience, heteroskedastic model *

	Whole Sample		S4 Leaver	
	Non-YT	*YT*	*Non-YT*	*YT*
Average	0.72	0.76	0.60	0.68
Advantaged	0.77	0.81	0.81	0.86
Disadvantaged	0.31	0.36	0.06	0.08

* *See notes to table 6.17a*

Taken over the whole SYPS employment sample, the heteroskedastic model results in a reduction of approximately one percent in the estimated impact for each of the stereotypical individuals described previously. The effect of participating in YT is seen from table 6.17b to increase the probability of being in employment by approximately 4 percentage points in the case of either an average or an advantaged school leaver. In the case of a disadvantaged leaver YT participation is associated with an increased employment probability of some 5 percentage points, two percent less than the uncorrected estimates. The effect of applying the heteroskedastic model is similar in the case of S4 leavers, with the advantage associated with participation

174 As pointed out by O'Higgins (op cit. p. 612), the heteroskedastic model as applied may pick up other forms of mis-specification and "thus what is called heteroskedasticity may, in reality, be more properly thought of as, for example, incorrect functional form."

in YT, in terms of expected employment probability, being marginally reduced. For an individual with an average set of characteristics the estimated employment impact of YT falls from 9 to 8 percentage points and from 7 to 5 percentage points for an advantaged leaver and, in the case of the disadvantaged S4 leaver, YT confers an advantage of just 2 percentage points. Application of the heteroskedastic model does not, however, materially change the conclusions drawn from the uncorrected Probit.

As in the earnings case, it should be noted that these results are again based on observations early in the young person's working life. Moreover, they are for a specific point in time and reveal little about the many different routes by which employment status is achieved at that point in time or how long particular individuals have spent in a particular. In other words, the Probit fails to take into account the potential of the training programme to place participants in a job sooner than non-participants (or would have otherwise been the case) and hence does not adequately describe the school-to-work transition state (Dolton, et al, 1994c, pp. 632-633). In order to assess employment probability over time we apply duration analysis to the employment sample drawing on the diary information collected in the SYPS. The diary entries were crosschecked with other dated survey questions to ensure consistency enabling the construction of a duration variable indicating the time, in months, and individual took to start on their first job after entering the labour market. Finally, following Dolton et al. (1994c, pp. 636-638), the time spent on YT was subtracted from the duration variable of participants. Missing data and inconsistency in responses contributed to a significant reduction in the size of the respective samples, reducing the full sample (all school leavers) to 987 cases and the S4 leavers sample to just 479 cases. Data constraints also prevent any reliable determination of the quality of the job as Dolton et al. (ibid. p. 637) attempted to do in differentiating between 'good' and 'bad' jobs. They did not, however, identify any significant difference when making this distinction as opposed to estimating the time to 'any' job.

Table 6.18 presents the estimates of the time to find a job based on Cox's Proportional Hazards Model (formally detailed in section 3.5.2, equations 3.36 to 3.38) for our two sample groups. Those respondents remaining in unemployment at the end of the period (October 1993) are effectively censored in this form of analysis. This semi-parametric method allows observed heterogeneity as represented by individual characteristics to be taken into account. The descriptive variables included in the Probit equations are retained excepting those with a time component, such as months of unemployment and time in training. A positive coefficient implies that the requisite variable increases the chance of finding a job at each point in time, and *vice versa* in the case of a negative coefficient. The partial likelihood estimates (based on equation 3.38 in chapter 3) for the full sample are reported in the first column of table 6.18 and indicate that YT participants are likely to start a job before non-participants especially if they also gain a qualification from the YT course.

Table 6.18: Cox Proportional Hazards Model estimates of duration to first job, using SYPS data*

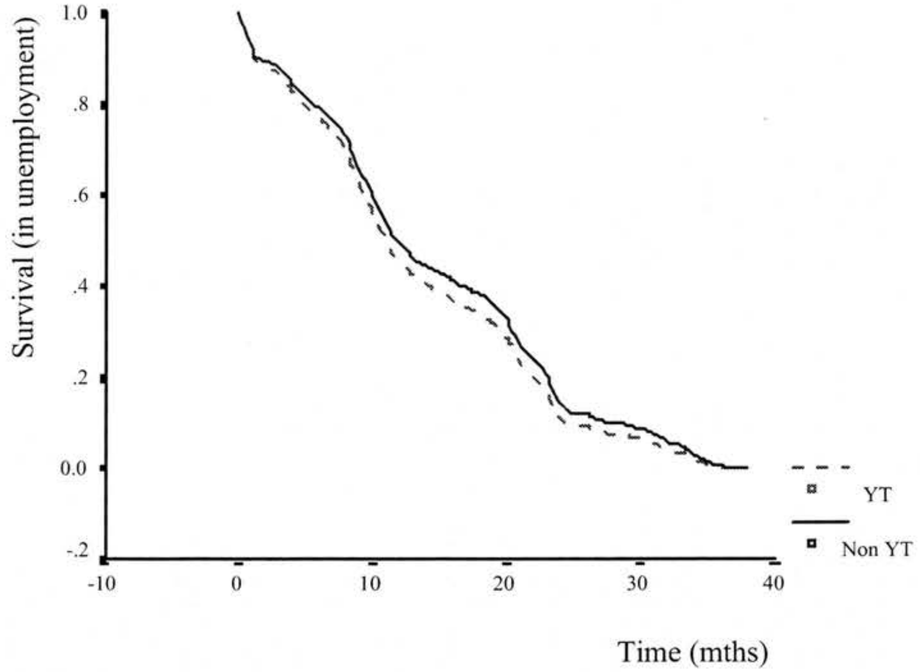
Descriptive variables	*Whole Sample*	*S4 Leavers*
Personal characteristics		
Female	-0.128 (1.81)	-0.262 (2.43)
Married	0.317 (2.46)	0.400 (2.18)
School experience		
Unqualified	-0.006 (0.04)	-0.141 (0.86)
1-3 O/S-levels	0.086 (0.72)	-0.020 (0.14)
4+ O/S-levels	0.001 (0.10)	-0.006 (0.41)
Highers	0.051 (0.37)	..
S4 leaver	1.182 (8.35)	..
S5 leaver	0.660 (6.06)	..
Habitual truant	-0.139 (0.91)	-0.044 (0.26)
Part time job while at school	0.148 (1.94)	0.191 (1.83)
Background characteristics		
Father unemployed	-0.200 (1.26)	0.035 (0.16)
Father in skilled/ manual occupation	0.207 (1.95)	0.234 (1.52)
Father in professional/ intermediate	0.211 (1.80)	0.320 (1.80)
Father's occupation not given	0.115 (0.88)	0.198 (1.12)
Mother in non-manual occupation	0.026 (0.37)	0.045 (0.41)
Labour market experience		
Local unemployment rate	-0.009 (0.86)	-0.019 (1.26)
Youth Training		
YT Experience	0.058 (1.30)	-0.250 (2.34)
YT Qualification	0.086 (0.65)	0.117 (0.73)
Chi-square	117.54	27.07
Log-likelihood	-5052.81	-2100.55
N (number of observations censored)	*987 (121)*	*479 (79)*

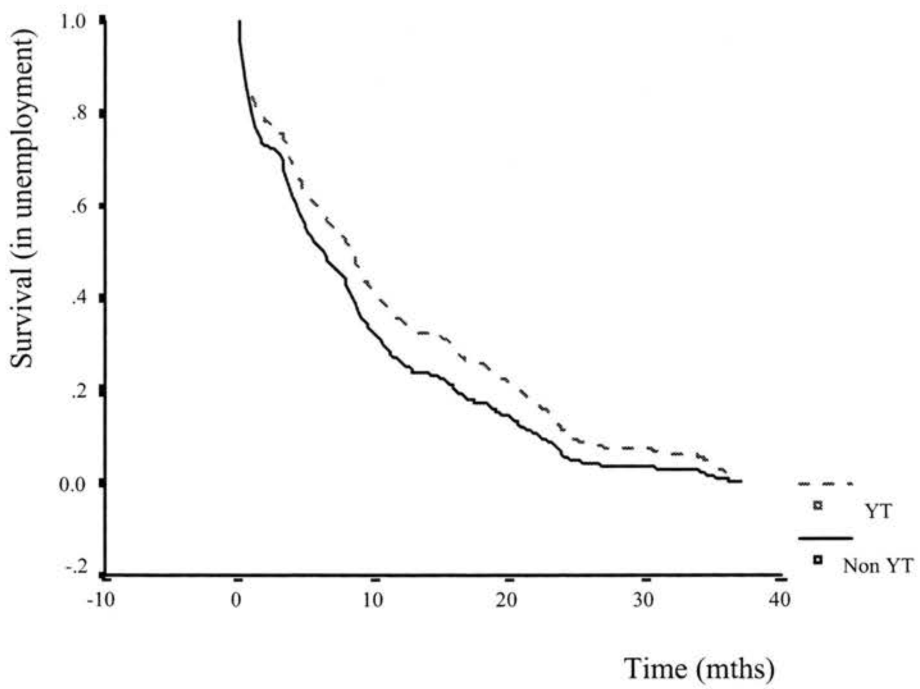
**coefficient (t-statistic)*

The effect of YT participation can be more readily observed in figure 6.2a, which shows the resultant plots of the survival function for the sample after the application of Proportional Hazard Model, by YT status. The plot clearly shows that individuals are moving into jobs over the period of the survey and that there is advantage accrued to YT participation (broken line), which reinforces the findings of the Probit analysis.

When only those who left school before Christmas 1990 are included in the analysis, the estimates of the Proportional Hazards Model, reported in the second column of table 6.18, indicate that the impact of YT is to reduce the probability of finding work at each point in time. This negative impact is partly ameliorated in the case where the participant gains a qualification from their involvement in YT. The plots in figure 6.2b show the extent of the disadvantage associated with YT participation (broken line) when no qualification is gained. The extended length of time taken to start their first job may, in part, explain the relatively poor impact of YT on the employment prospects of disadvantaged S4 leaver observed in the Probit analysis. As illustrated in Dolton et al. (1994c) and Mealli, Pudney and Thomas (1996), among others, duration analysis can be extended in an impressive variety of directions, including the explicit modelling of sample selection and the inclusion of time varying covariates (such as the local unemployment rate). However, given the limitations to the SYPS sub-sample discussed previously, these are not developed here.

Figure 6.2a: Survival plots by YT status, all school leavers





6.6: Chapter Summary

This chapter has examined a number of aspects of the evaluation process using the Scottish Young Person's Survey, which concentrates on the immediate school leaving period of respondents. The advantage of this type of dataset for programme evaluation rests with the breadth of information on both participants and non-participants. The data is quite rich with detail on individuals' schooling experience, family background and the school-to-work transition, at least with respect to the first job or programme. Information at the time of the follow-up sweep also appears to be of good quality but the intermediate period, which largely rests on the respondent's recall, is limited in its ability to capture the dynamic aspects of the labour market. It is also limited by its duration in only covering the period in which respondents are approximately 16 years of age through to when they are approximately 19 years of age. Nevertheless, this chapter has illustrated that the majority of the selection models outlined in chapter three can be effectively applied using this dataset.

Concentrating first on the earnings impact of YT, sensitivity analysis clearly illustrated that model choice is likely to have contributed towards much of the variability observed in impacts estimates in the UK. The variability in estimates across the different models also suggests that selectivity bias may be a significant factor in the SYPS data even though it was not identified in the form of a *t-test* on the selectivity variable (Lambda) in the standard Heckman two-step model. Inconsistency in the direction of impact across the models constrains the interpretation of the results, however, for the majority of participants the early returns to training would appear to be negative, which is consistent with the evaluation literature on modern government assisted training programmes. On the other hand, for participants who remain on a programme for less than 12 months but still attain a qualification the impact was unambiguously positive. The results were also promising for individuals fitting the disadvantaged stereotype, but in both cases these benefited only a small portion of the sample and rested on estimates that were relatively weak statistically. The analysis would suggest that it would be good practice to apply more than one model when estimating the returns to training. It has also highlighted the importance of taking into account the structure of the programme, in this case the variables representing the duration of training tended to have a significant effect on the outcome estimates.

In terms of the employment impact of YT, the analysis revolved around the standard Probit model which was also estimated using a correction for heteroskedasticity. Application of the heteroskedastic model resulted in marginally smaller estimates, however the findings were consistent with the conclusions drawn from the standard model. Again, consistent with the existing literature, it is observed that the effect of YT is quite small in comparison with the personal circumstances and characteristics that differentiate the disadvantaged from the average and advantaged stereotypes. The analysis is reinforced by duration analysis that shows that participants are consistently moving into jobs over the period of the survey at a superior rate to non-participants although this cannot make up for factors that contribute to disadvantage such as leaving school at the earliest possible opportunity.

Both the earnings and employment impacts are identified as proxy measures for the general aims of the programme, which have tended to be defined in relatively broad-brush terms. If the programme could be shown to be effective for both outcomes, even in the short term, it would suggest that the programme is meeting most of its aims. In the case of early returns to YT participation in Scotland, however, this would appear to be the case for only a very small proportion of participants. This type of analysis nonetheless has helped identify what types of individuals are most likely to benefit from the programme. The next chapter summarises the findings of the thesis, identifies potential applications of this research and maps out directions for future research. It also includes possible policy implications arising from the analysis and a discussion on the design and collection of non-experimental data necessary for more effective training evaluation.

Chapter 7: Conclusions

On leaving school young people who do not enter higher education or training continue to enter an uncertain labour market. They face relatively high unemployment rates and are more likely to be laid off in economic downswings. The costs of unemployment are well documented, but these are likely to be amplified for the young in terms of undermining their motivation and skill development as well as their social and fiscal well-being. For the economy as a whole, the loss of such skill development is heightened as countries strive to maintain or improve their international competitiveness. Such concerns together with rapidly changing patterns in labour demand and an ageing workforce combine to ensure that the youth labour market remains an area of high priority in most modern economies. The response in the majority of OECD countries has been to increase intervention in the school-to-work transition with the aim of minimising youth unemployment and encouraging skill development. In the case of the United Kingdom this has taken the shape of large-scale youth training programmes such as the YTS and the New Deal. These factors and the requisite responses were detailed in Chapter 2 which provided the necessary context for the research conducted in this thesis by assessing the state of the youth labour market in a selection of OECD countries. Focusing on the unemployment rate, the labour force participation rate and the employment/population ratio it was found that, in general, the state of the youth labour markets was less than desirable and exacerbated in times of economic recession such as the 1970s. Moreover the plight of some groups, particularly young men, was even worse. The seriousness of this issue ensured that it remained high on the policy agenda reflected in the growth of active labour market programmes, including youth training.

In order to determine the success of labour market interventions accurate on-going evaluation is required. However, because it is not possible to simultaneously observe the subject in both the treated and untreated state, evaluations of programme impact are likely to be subject to bias. The primary source of this bias arises from sample selection, which occurs when a rule other than simple random sampling determines the sample gathered from the underlying population of interest. For example, in the case of training, observed outcomes are not randomly assigned across the population but are a result of individual choices, circumstances and attributes. In economics two key approaches, experimental and non-experimental, are adopted to help resolve any sample selection bias. The debate between the relative merits of these two approaches was used as a framework, in chapter 3, to detail the theoretical developments in the literature. This included identifying the necessary ingredients for efficient impact analysis in the field of youth training programme evaluation. This principally involves the construction of a realistic counterfactual either through the use of an experimental or statistical control group, as failure to do so renders estimates unreliable. It was observed that experimental methods appear to offer the most promising solution to the evaluation problem but the implementation of a successful social experiment faces a number of restrictive limitations. These range from the practical difficulties arising from randomisation and substitution bias through to the ethical conflict of having to deny eligible applicants places on programmes. Moreover, at the current time, social experiments are not a feasible political option for the evaluation of training programmes in the UK or many other modern economies and hence the availability of experimental data is severely constrained.

The thesis then focuses on non-experimental methods and the remainder of chapter 3 examined the main theoretical advances that have been made in this field detailing the different non-experimental estimators and examining various methods that might be used for choosing between them. As observed above, the fundamental finding of theoretical investigation into non-experimental methods for training programme evaluation is that there is a high likelihood that OLS estimates will be biased because of sample selection. While a wide variety of econometric models have been proposed to correct for this bias it is unlikely that any single method will always solve the problem, as all of the non-experimental models require at least one minimal identifying assumption. Because minimal identifying assumptions cannot be tested with data, model choice can hence prove to be complex and controversial. Further, a number of factors need to be considered when choosing an appropriate model, these range from the nature of the training programme to the availability of appropriate data, however the actual objectives of the programme may be poorly defined impacting on both selection of outcome measures and model choice. Moreover, the frequent absence of any formal structure and consequential variability in the delivery, accreditation and setting of training standards not only clouds identification of any specific programme objectives but also hampers the definition of what might actually constitute training. The definition of training is further complicated through recent institutional reforms which have witnessed a blurring of the boundary between education and training. Consequently, in terms of outcomes, it is often difficult to disentangle the impact of training from other influences, exacerbated by extensive overlap in public and private sector provision, all of which complicate model choice and design.

Some of these issues were addressed in chapter 4 which presented an overview of recent institutional developments associated with the provision of youth training in the UK and Australia. It also critically reviewed the relevant empirical literature on youth training programme evaluations by various outcomes of interest and presented an overview of relevant datasets. Focusing on non-experimental studies in the UK it was shown that, even though the studies evaluated similar programmes and used seemingly similar datasets, a wide range of impact estimates were reported, particularly when measured in terms of the probability of gaining employment. This observation was consistent with the findings of earlier surveys of evaluation studies in both the UK and the USA, suggesting that such variation may be inherent in any pool of non-experimental evaluations. However, a summary of some recent experimental evaluation results conducted on training schemes in the USA also showed significant variation in reported estimates illustrating that the problem might not be confined to the non-experimental approach.

The literature survey showed that a wide range of economic and statistical specifications has been employed in the UK evaluations, partly in response to the constantly evolving nature of the youth training programme(s). Specifications have also changed in response to the identification of different sources of bias, ranging from those inherent in the datasets, to those in the choice of model variables and structure. This has led to evaluations being somewhat piecemeal and varied, in addition to being applied to different surveys and survey sweeps making it difficult to draw any firm conclusions regarding the practical importance or relevance of some of these innovations. Such issues were addressed in chapter 6, where I employed sensitivity analysis to help reconcile some of the observed variation in impact estimates; this involved the application of an array of feasible specifications to a common database hence controlling for influences due to data, programme and other time related factors.

A different approach to the UK evaluation literature was adopted in chapter 5 which presented the results of a meta-analysis conducted on the studies surveyed in the previous chapter. The motivation behind this approach was to provide a more objective identification of the sources of variance in estimates across similar studies and to achieve greater accuracy than that obtainable from a single study. This is based on the premise that greater accuracy is gained through the statistical analysis of the body of research being reviewed, in which the summary statistics from the original studies are used as the data-points for the meta-analysis. After presenting an overview of the development of meta-analysis in the social sciences and reviewing a selection of existing meta-analyses within the economics literature, the chapter revisited the UK studies investigating a range of potential sources of variation including the dataset, author, allowance for sample selection, and other specification factors. The meta-analysis found that much of the observed between-study variation in estimates is highly likely to be due to the model employed. For example, sample selection models were found to result in estimates substantially different from the cross-study mean when used to estimate the impact in terms of either employment probability or relative post-training earnings. Because of the limited sample size, the results were only indicative but they clearly illustrate the potential of meta-analysis to enhance future literature reviews of empirical studies. It is also possible that the development and application of a formal procedure for detecting statistically significant differences across a group of related studies can also aid model development and the transfer of knowledge to similar research questions.

The available pool of evaluation studies in the UK was, however, too limited to allow meta-analysis to establish the influence of specific (selection) model types on impact estimates. As observed above, this issue was hence investigated further in chapter 6 which used sensitivity analysis to examine between-study variation arising from methodological choices made by the analyst. This approach hence controlled for any between-study variation that may arise from the use of different programmes and/or datasets. It is then possible to investigate the influence of the various assumptions, inherent in the application of different evaluation models, in terms of their impact on estimates of the returns to training. The focus of the sensitivity analysis in chapter 6 was on the post-training earnings of young people in Scotland, drawing on data from the Scottish Young Person's Survey (SYPS). Longitudinal surveys, such as the SYPS, provide detailed information on the immediate post-school experience of respondents - both training programme participants and non-participants. The data proved to be quite rich with detail on individuals' schooling experience, family background and the school-to-work transition, at least with respect to the first job or programme. However, the nature of the data means that we can only gain historic insights into the effectiveness of government intervention in the school-to-work transition because training and education curriculum and structures are constantly evolving. This study hence reported the outcome of training structures that were in place in the early 1990s.

The impact of youth training (YT) is reported in terms of both post-programme earnings and on the probability of subsequently gaining employment. When applied to a variety of models commonly used to assess the earnings impact of YT, the key findings of the sensitivity analysis were, first, that model choice is likely to have contributed towards much of the variability observed in impact estimates in the UK. Second, this variability in estimates suggests that selectivity bias is a significant factor in the SYPS data even though it was not identified in the form of a *t-test* on the selectivity variable (Lambda) in the standard Heckman two-step model. Third, for the majority of participants the early returns to training, in terms of post-programme earnings impact, would appear to be negative, which is consistent with the evaluation literature on modern government assisted training programmes. The observed inconsistency in the direction of impact across the models necessitates the caveat to the latter that the results are only indicative. The analysis also involved the application of the analysis to different stereotypical profiles based on advantage, in terms of family and social background, and school leaving status. A positive earnings impact was observed for those individuals fitting the disadvantaged stereotype and, in all cases, the models consistently estimated a positive impact for those participants who remained on a programme for less than 12 months but still attained a formal qualification. The second of these findings initially seemed implausible but subsequent investigations of programmes on offer at the time uncovered a number of short courses that led to qualifications, particularly in the information technology and hospitality industries. While these findings only benefited a small portion of trainees in this cohort of the SYPS, they both have interesting policy implications in terms of both targeting (and distributional issues) and programme design. The short course and qualification scenario is also in keeping with the flexibility and design of delivery in the latest evolution of training programme in the UK, the New Deal. In general, the results of the sensitivity analysis would suggest that it would be good practice to apply more than one model when estimating the returns to training in terms of the impact on post-programme earnings. It has also highlighted the importance of taking into account the structure of the programme, in this case the variables representing the duration of training tended to have a significant effect on the outcome estimates.

In terms of the post-programme employment impact of YT, it was observed that the effect of YT, although positive, is quite small in comparison with the personal circumstances and characteristics that differentiate the disadvantaged from the average and advantaged stereotypes. This finding was consistent for analysis based both on the Probit model and duration models. In either the earnings or employment approach, the impact measures used are identified as only proxy measures for the general aims of the programme. As observed above, this is largely because the aims and objectives of the programme have been broadly defined and clouded by variability in programme funding, structure, delivery, accreditation and setting of training standards.

The current approach to evaluation research can be easily summarised in a simple pay-off matrix, illustrated in figure 7.1 below, which incorporates the two key outcome measures used to assess the impact of youth training schemes, earnings and employment. Clearly if both measures result in a positive outcome the intervention can be considered successful, with an equally unambiguous conclusion when both outcomes are negative. However, the more interesting cases arise when only one of the outcomes is positive. In the USA post-programme earnings are considered to be illustrative of programme success, in effect addressing market failure, and may be given greater weighting, however this is clearly not the only criterion of programme success in the UK. For example, the analysis in chapter 6 showed that only a small group benefited from participation in terms of employment but if distributional issues are important to the policy makers then, clearly this may be given more weight. Equally, factors such as equality of opportunity, alleviating unemployment black-spots and long-term welfare dependency may be given greater emphasis, together with alternative outcomes such as qualifications gained.

Figure 7.1: Outcome matrix

Employment \ Earnings	*Positive*	*Negative*
Positive	+	0
Negative	0	-

This again illustrates the need for a programme to have clearly defined aims and objectives. This, in turn, impacts on the data requirements to ensure that both the training and outcome can be clearly identified and measured. To control for factors such as variability in delivery of training, together with differing social or environmental characteristics, it is important that information on both participants and non-participants be collected from the same region. This would open the door to the application of more developed matching techniques, which would also benefit from greater detail on the supply of training, such as the training cost, assessment methods (in-house or external evaluation, practical or theoretical, etc.) and eligibility, along with supply related factors like firm size and locality. Some of this information could be more readily compiled if its collection by the training provider was made a condition of any government subsidy on training provision. Linking to the recipient data could then be made through a unique identification code or number, such as the national insurance number, which would still allow for the necessary privacy of both participants and non-participants. Such broad compilation of data would not only improve the construction of control groups but also more readily allow the application of recently developed techniques such as quantile regressions. Naturally, all the evaluation techniques would benefit from greater duration in the monitoring period and requisite data collection.

In summary, this thesis has addressed some of the inconsistencies observed in existing non-experimental impact estimates in the United Kingdom by applying a range of evaluation methods to a common dataset and conducting a sensitivity analysis. It can be seen that much of the observed variability is due to the model employed. Given that the different models rest on different assumptions the analysis does not, however, help us determine which is the appropriate model for assessing this type of programme. It does suggest that good practice would involve the application of more than one model and that consistency in the direction of impact measures guide any policy recommendations. To answer the question pertaining to the accuracy of the impact measures, it is clear that the efficacy of the non-experimental models can only be fully ascertained with better data. Better data, however, need not imply that experimental data is the only alternative, as this too is subject to limitations.

Bibliography

Ainley, J., J. Malley, and S. Lamb. (1997) "Thematic Review of the Transition from Initial Education to Working Life: Australia, background report." OECD, Paris.

Aldrich, J. H. and F. D. Nelson. (1984) "Linear Probability, Logit, and Probit Models." in *Quantitative Applications in the Social Sciences, 07-045*, vol. 45, edited by J. L. Sullivan. Beverly Hills: Sage Publications.

Amemiya, T. (1973) "Regression analysis when the dependent variable is truncated normal." *Econometrica, 41*, 997-1016.

Amemiya, T. (1981) "Qualitative response models: A survey." *Journal of Economic Literature, 19*, 1483-1536.

Andrews, M. and S. Bradley. (1997) "Modelling the transition from school and the demand for training in the United Kingdom." *Economica, 64*, 387-413.

Arulampalam, W., J. Hartog, T. MacCurdy, and J. Theeuwes. (1997) "Replication and re-analysis." *Labour Economics, 4*, 99-105.

Ashenfelter, O. (1978) "Estimating the effect of training programs on earnings." *The Review of Economics and Statistics, 60*, 47-57.

Ashenfelter, O. (1987) "The case for evaluating training programs with randomized trials." *Economics of Education Review, 6*, 333-338.

Ashenfelter, O. (1994) "How convincing is the evidence linking education and income." *Labour Economics and Productivity, 6*, 274-285

Ashenfelter, O. and D. Card. (1985) "Using the longitudinal structure of earnings to estimate the effect of training programs." *The Review of Economics and Statistics* :648-660.

Assmus, G., J. U. Farley, and D. R. Lehmann. (1984) "How advertising affects sales: Meta-analysis of econometric results." *Journal of Marketing Research, 21*, 65-74.

Australian Bureau of Statistics [ABS]. (various) "Ausstats." ABS Website, http://www.abs.gov.au/ausstats.

ABS. (various) *Year Book Australia*. Canberra: ABS.

Australian National Training Authority [ANTA]. (1997) *Stocktake of equity reports and literature in vocational education and training*. Canberra: ANTA.

ANTA. (1999) "ABC of VET." ANTA website, http://www.anta.gov.au /anta_prod/ABC.

Baaijens, S. R. and P. Nijkamp. (1997) "Explanatory meta-analysis of tourist income multipliers." *Tinbergen Institute Discussion Papers*.

Baker, M. A. (1990) "The effect of training on the earnings of young males: An analysis of the National Child Development Study." Masters Thesis, University of Warwick, Warwick.

Baker, M. A. (1994) "Training down under: An overview of the Australian experience." *International Journal of Manpower Studies, 15*, 70-88.

Bangert-Drowns, R. L. (1987) "Review of developments in meta-analytic method." pp. 655-666 in *Evaluation Studies Review Annual*, edited by W. R. Shadish and C. S. Reichardt, Beverly Hills: Sage.

Barnow, B. S. (1987) "The impact of CETA programs on earnings: A review of the literature." *Journal of Human Resources, 22*, 157-193.

Barnow, B. S., G. G. Cain, and A. S. Goldberger. (1980) "Issues in the Analysis of Selectivity Bias." pp. 43-59 in *Evaluation Studies Review Annual*, edited by E. W. Stromsdorfer and G. Garkas. Beverly Hills: Sage.

Barwell, R. (2000) "Age structure and the UK unemployment rate." *Bank of England Working Paper, 124*.

Bassi, L. J. (1984) "Estimating the effect of training programs with non-random selection." *The Review of Economics and Statistics, 66*, 36-43.

Becker, G. S. (1962) "Investment in human capital: A theoretical analysis." *Journal of Political Economy, 70*, 9-49.

Begg, I. G., A. P. Blake, and B. M. Deakin. (1991) "YTS and the labour market." *British Journal of Industrial Relations, 29*, 223-236.

Bell, S. H., L. R. Orr, J. D. Blomquist, and G. G. Cain. (1995) *Program Applicants as a Comparison Group in Evaluating Training Programs*. Michigan: W.E. Upjohn Institute.

Bera, A. K., C. M. Jarque, and L.-F. Lee. (1984) "Testing the normality assumption in limited dependent variable models." *International Economic Review, 25*, 563-578.

Bergh (van den), J. C. J. M., K. J. Button, P. Nijkamp, and G. C. Pepping. (1997) "Meta-Analysis in Environmental Economics." in *Economy and Environment*, vol. 12. London: Kluwer Academic Publishers.

Berndt, E. R. (1991) *The Practice of Econometrics: Classic and Contemporary*: Addison-Wesley.

Bjorklund, A. and R. Moffitt. (1987) "The estimation of wage gains and welfare gains in self-selection models." *The Review of Economics and Statistics* :42-49.

Bjorklund, A. and H. Regner. (1996) "Experimental Evaluation of European Labour Market Policy." pp. 89-114 in *International Handbook of Labour Market Policy and Evaluation*, edited by G. Schmid, J. O'Reilly, and K. Schomann. London: Edward Elgar.

Blanchflower, D. G. and R. Freeman. (1996) "Youth employment and joblessness in advanced countries." in *NBER Conference on Youth* 12-14 December 1996.

Blanchflower, D. G. and R. Freeman. (1999) "The declining economic status of young workers in OECD countries." in *NBER Conference on Youth* 19 March 1999.

Blanchflower, D. G. and L. Lynch, M. (1995) "Training At Work: A Comparison of U.S. and British Youths." pp. 233-260 in *International Comparisons of Private Sector Training*, edited by L. M. Lunch. Chicago: University of Chicago Press.

Bloch, F. (1979) *Evaluating Manpower Training Programs*. Supplement 1 to *Research in Labor Economics*, Greenwich: JAI Press.

Blundell, R., L. Dearden, C. Meghir, and B. Sianesi. (1999) "Human capital investment: the returns from education and training to the individual, the firm and the economy." *Fiscal Studies, 20*, 1-23.

Bradley, S. (1995) "The Youth Training Scheme: A critical review of the evaluation literature." *International Journal of Manpower, 16*, 30-56.

Bradley, S. (1995) "An empirical analysis of private sector training." *University of Lancaster Discussion Papers in Economics* (8/95).

Breen, R. (1988) "The work experience programme in Ireland." *International Labour Review, 127*, 429-444.

Breen, R. (1991) *Education, Employment and Training in the Youth Labour Market*, vol. 152. Dublin: The Economic and Social Research Institute.

Breen, R. (1996) "Regression Models: Censored, Sample Selected, or Truncated Data." in *Quantitative Applications in the Social Sciences, 07-111*, vol. 111, edited by M. S. Lewis-Beck. Thousand Oaks: Sage Publications.

Breusch, T. S. and A. R. Pagan. (1979) "A simple test for heteroscedasticity and random coefficient variation." *Econometrica, 47,* 1287-1294.

Buchinsky, M. (1994) "Changes in the U.S. wage structure 1963-1987: application of quantile regression." *Econometrica, 62*, 405-458.

Buchinsky, M. (1997) "Recent advances in quantile regression models: A practical guideline for empirical research." *The Journal of Human Resources, 33*, 88-126.

Buechtemann, C. F., J. Schupp, and D. Soloff. (1993) "Roads to work: school-to-work transition patterns in Germany and the United States." *Industrial Relations Journal, 24*, 97-111.

Burtless, G. (1995) "The case for randomised field trials in economic and policy research." *Journal of Economic Perspectives, 9*, 63-84.

Burtless, G. and L. R. Orr. (1986) "Are classical experiments needed for manpower policy." *Journal of Human Resources, 21*, 606-639.

Buxton, T., P. Chapman, and P. Temple. (1998) "Britain's Economic Performance." London: Routledge.

Card, D. and A. B. Krueger. (1995) *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton: Princeton University Press.

Card, D. and D. Sullivan. (1988) "Measuring the effect of subsidised training programs on movements in and out of employment." *Econometrica, 56*, 497-530.

Centre for Educational Sociology. (various) *Scottish Young Person's Survey Documentation*. Edinburgh: University of Edinburgh.

Chamberlain, G. (1994) "Quantile Regression, Censoring, and the Structure of Wages." in *Advances in Econometrics, Sixth World Congress*, vol. 1, edited by C. Sims, A: Cambridge University Press.

Chapman, B. J. and P. J. Stemp. (1992) "Government intervention in the provision of on-the-job training." *Australian Economic Papers, 31*, 354-368.

Chapman, P. G. and M. J. Tooze. (1987) "Some economic implications of the Youth Training Scheme." *The Royal Bank of Scotland Review*: 14-23.

Chesher, A. and M. Irish. (1987) "Residual analysis in the grouped and censored normal linear model." *Journal of Econometrics, 34*, 33-61.

Clark, K. B. and L. H. Summers. (1982) "The Dynamics of Youth Unemployment." pp. 199-235 in *The Youth Labor Market Problem: Its Nature, Causes and Consequences*, edited by R. B. Freeman and D. A. Wise. Chicago: University of Chicago Press.

Cochran, W. G. (1937) "Problems arising in the analysis of a series of similar experiments." *Journal of the Royal Statistical Society* (Supplement 4), 102-118.

Commonwealth Department of Employment Education and Training. (1996) *Australian Youth Survey 1989-1994*. Canberra : Social Science Data Archives: The Australian National University (*computer file*).

Daly, A. (1993) "The determinants of employment for aboriginal people." *Australian Economic Papers* (June), 134-151.

Davidson, R. and J. G. MacKinnon. (1984) "Convenient specification tests for Logit and Probit models." *Journal of Econometrics, 25*, 241-262.

Davidson, R. and J. G. MacKinnon. (1993) *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Davies, P. and N. Adnett. (1999) "Quasi-market reforms and vocational schooling in England." *Journal of Education and Work*. 12(2), 139-154.

Deakin, B. M. (1996) *The Youth Labour Market in Britain: The Role of Intervention*. Cambridge.

Dearden, L. (1995) "Education, Training and Earnings in Australia and Britain", PhD. Thesis, University College London, London.

Dearden, L. and A. J. Heath. (1996) "Income support and staying in school: What can we learn from Australia's AUSTUDY experiment?" *Fiscal Studies, 17*, 1-30.

Dehejia, R. H. and S. Wahba. (1998) "Causal effects in non-experimental studies: Re-evaluating the evaluation of training programmes." *NBER Working Paper* (No. 6586).

Department for Education and Employment [DfEE]. (1994) *The Youth Cohort Study: An Information Pack*: DfEE.

DfEE. (1998) "Education and Labour Market Statistics of Young people Aged 16-18 in England: 1992-1997." DfEE, London.

DfEE. (1999) "Participation in Education and Training by 16-18 Year Olds in England: 1988 to 1998." DfEE, London.

DfEE. (various) *New Earnings Survey*. London: HMSO.

DfEE. (various) *Education and Training Statistics for the United Kingdom*. London: HMSO.

DfEE. (various) "News Updates." DfEE Website, http://www.dfee.gov.uk/news.

Department for Education Training and Youth Affairs. (various) *Annual Report*. Canberra: DETYA.

Department for Trade and Industry. (1998) *Our Competitive Future: Building the Knowledge Drive Economy,* Cm 4176, The Stationary Office, London

Dewalk, W. G., J. Thursby, G, and R. G. Anderson. (1986) "Replication in empirical economics: The Journal of Money, Credit and Banking project." *The American Economic Review, 76*, 587-603.

Dickinson, K. P., T. R. Johnson, and R. W. West. (1987) "An analysis of the sensitivity of quasi-experiemental net impact estimates of CETA programs." *Evaluation Review, 11*, 452-472.

DiNardo, J., Fortin, N. and Lemieux, T. (1996) "Labor Market Institutions and the Distribution of Wages 1973-1992: A Semiparametric Approach". *Econometrica* Vol 64, 1001-1044.

Dolton, P. J. (1993) "The Economics of youth training in Britain." *The Economic Journal, 103*, 1261-1278.

Dolton, P. J. and G. H. Makepeace. (1986) "Sample selection and male-female earnings differentials in the graduate labour market." *Oxford Economic Papers, 38*, 317-341.

Dolton, P. J. and G. H. Makepeace. (1987) "Interpreting sample selection effects." *Economics Letters, 24*, 373-379.

Dolton, P. J., G. H. Makepeace, and J. G. Treble. (1994a) "Measuring the Effects of Training in the Youth Cohort Study." pp. 195-230 in *The Market for Training*, edited by R. McNabb and K. Whitfeild. Aldershot: Avebury.

Dolton, P. J., G. H. Makepeace, and J. G. Treble. (1994b) "The wage effect of YTS: Evidence from YCS." *Scottish Journal of Political Economy, 41*, 444-453.

Dolton, P. J., G. H. Makepeace, and J. G. Treble. (1994c) "The Youth Training Scheme and the school-to-work transition." *Oxford Economic Papers, 46*, 629-657.

Dolton, P. J., G. H. Makepeace, and J. G. Treble. (1995) "Public and Private-Sector Training of Young People in Britain." in *International Comparisons of Private Sector Training*, edited by L. M. Lynch. Chicago: University of Chicago Press.

Dolton, P. J. and D. O'Neill. (1996) "Unemployment duration and the restart effect: some experimental evidence." *The Economic Journal 106*, 387-400.

Donald, S. G. (1995) "Two-step estimation of heteroskedastic sample selection models." *Journal of Econometrics, 65*, 347-380.

Doolittle, F. and L. Traeger. (1990) *Implementing the National JTPA Study*: Manpower Demonstration Research Corporation.

Doucouliagos, C. (1997) "The aggregate demand for labour in Australia: A meta-analysis." *Australian Economic Papers,* (December) 224-242.

Elias, P. (1995) "Training in Britain: The quiet revolution." *The Australian Economic Review*: 109-115.

Fay, R. G. (1996) "Enhancing the effectiveness of active labour market policies: Evidence from OECD countries." *Labour Market and Social Policy Occasional Papers* (No. 18, OECD, Paris).

Favennec-Hery, F. (1996) "Work and training a blurring of the edges." *International Labour Review, 135 (6)*, 665-673.

Finifter, D. H. (1987) "An approach to estimating net earnings impact of federally subsidized employment and training programs." *Evaluation Review, 11*, 528-547.

Fisher, R. A. (1932) *Statistical Methods for Research Workers*. London: Oliver and Boyd.

Frakcr, T. and R. Maynard. (1987) "The adequacy of comparison group designs for evaluations of employment-related programs." *Journal of Human Resources, 22*, 194-227.

Freeman, R. B. (1996) "Why do so many young American men commit crime and what might we do about it?" *Journal of Economic Perspectives, 10 (1)*, 25-42.

Friedlander, D., D. H. Greenberg, and P. K. Robins. (1997) "Evaluating government training programs for the economically disadvantaged." *Journal of Economic Literature, 35*, 1809-1855.

Friedlander, D. and P. K. Robins. (1995) "Evaluating program evaluations: New evidence on commonly used non-experimental methods." *American Economic Review, 85*, 923-937.

Fry, T. R. L., R. D. Brooks, B. R. Comley, and J. Zhang. (1993) "Economics motivations for limited dependent and qualitative variable models." *The Economic Record, 69*, 193-205.

Guade, J. (1997) "L'Insertion des Jeunes et les Politiques d'Emploi Formation" *ILO Employment and Training Papers (1)* Geneva

Glass, G. V. (1976) "Primary, secondary, and meta-analysis of research." *Educational Researcher, 5*, 3-8.

Glass, G. V. (1977) "Integrating findings: The meta-analysis of research." *Review of Research in Education, 5*, 351-379.

Glass, G. V. and M. Smith. (1979) "Meta-analysis of research on class size and achievement." *Educational Evaluation and Policy Analysis, 1*, 2-16.

Glass, G. V., B. McGaw and M. L. Smith. (1981) *Meta Analysis in Social Research*. Beverly Hills: Sage Publications.

Goetz, S. J. and F. V. Glass. (1995) "Markets, transaction costs, and selectivity models in economic development." pp. 383-399 in *Prices, Products, and People: Analyzing Agricultural Markets in Developing Countries*, edited by G. J. Scott and Reinner. London.

Goldberger, A. S. (1981) "Linear regression after selection." *Journal of Econometrics, 15*, 357-366.

Goldstein, H. (1995) *Multilevel Statistical Models*. London: Edward Arnold.

Gospel, H. (1994) "The survival of apprenticeship training: A British, American and Australian comparison." *British Journal of Industrial Relations, 32*, 505-522.

Green, F. (1993) "The determinants of training of male and female employees in Britain." *Oxford Bulletin of Economics and Statistics, 55*, 103-122.

Green, F., M. Hoskins, and S. Montgomery. (1996) "The effects of company training, further education and the Youth Training Scheme on the earnings of young employees." *Oxford Bulletin of Economics and Statistics, 58*, 469-488.

Greenaway, D. (1993) "Economic aspects of youth training in industrialsed countries: editorial note." *The Economic Journal, 103*, 1259-1260.

Greene, W. H. (1981) "Sample selection bias as a specification error: comment." *Econometrica, 49*, 795-798.

Greene, W. H. (1993) *Econometric Analysis*. New York: Macmillan Publishing Company.

Greene, W. H. (1995) *Limdep Version 7.0: User's Manual*. New York: Econometric Software.

Grey, K., W. Beswick, C. O'Brien, and D. Ray. (1999) "Traineeship non-completion." DETYA, REB 1/99, Canberra.

Gronau, R. (1974) "Wage comparisons - a selectivity bias." *Journal of Political Economy, 82*, 1119-1143.

Grubb, D. (1994) "Direct and Indirect effects of Active Labour Market Policies in OECD Countries." pp. 181-213 in *The UK Labour Market: Comparative Aspects and Institutional Developments*, edited by R. Barrell. Cambridge: Cambridge University Press.

Grubb, D. (1995) "Evaluating job training programs in the United States: evidence and Explanations" *Training Policy Studies no. 17*, Training Policies Branch, ILO, Geneva, (also available as NCRVE paper no. MDS-1047)

Hall, W. (1995) "The national training reform agenda." *The Australian Economic Review* (2nd Quarter).

Ham, J. and R. J. LaLonde. (1994) "Looking into the black box: Using experimental data to find out how training works." pp. 231-241 in *The Market for Training*, edited by R. McNabb and K. Whitfield. Aldershot: Avebury.

Haveman, R. H. (1986) "Social experimentation and social experimentation." *Journal of Human Resources, 21*, 586-605.

Heath, A. and S. Y. Cheung. (1999) "Education and Occupation in Britain." in *From School-to-Work: A Comparative Study of Educational Qualifications and Occupational Destination*, edited by Y. Shavit and W. Muller.

Heckman, J. J. (1974) "Shadow prices, market wages, and labor supply." *Econometrica, 42*, 679-694.

Heckman, J. J. (1978) "Dummy endogenous variables in a simultaneous equation system." *Econometrica, 46*, 931-961.

Heckman, J. J. (1979) "Sample selection bias as a specification error." *Econometrica, 47*, 153-161.

Heckman, J. J. (1980) "Addendum to 'Sample selection bias as a specification error'." pp. 70-74 in *Evaluation Studies Review Annual*, edited by E. W. Stromsdorfer and G. Garkas. Beverly Hills: Sage.

Heckman, J. J. (1990) "Varieties of selection bias." *American Economic Review, 80*, 313-318.

Heckman, J. J. (1992) "Randomization and Social Policy Evaluation." pp. 201-230 in *Evaluating Welfare and Training Programs*, edited by C. F. Manski and I. Garfinkel. Massachusetts: Harvard University Press.

Heckman, J. J. (1998) "What should be our human capital investment policy." *Fiscal Studies, 19*, 103-119.

Heckman, J. J. and B. Honore. (1990) "The empirical content of the Roy model." *Econometrica, 58*, 1121-1149.

Heckman, J. J. and V. J. Hotz. (1989) "Choosing among alternative non-experimental methods for estimating the impact of social progams: the case of manpower training." *Journal of American Statistical Association, 84*, 862-877.

Heckman, J. J., V. J. Hotz, and M. Dabos. (1987) "Do we need experimental data to evaluate the impact of manpower training on earnings?" *Evaluation Review, 11*, 395-427.

Heckman, J. J. and R. Robb. (1986) "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes." pp. 63-113 in *Drawing Inferences from Self-Selected Samples*, edited by H. Wainer and Springer-Verlag. New York.

Heckman, J. J. and G. Singer. (1985) "Longitudinal Analysis of Labor Market Data." in *Econometric Society Monographs*, vol. 10: Cambridge University Press.

Heckman, J. J. and J. A. Smith. (1995) "Assessing the case for social experiments." *Journal of Economic Perspectives, 9*, 85-110.

Heckman, J. J. and J. A. Smith. (1996) "Experimental and Non-experimental Evaluation." pp. 37-88 in *International Handbook of Labour Market Policy and Evaluation*, edited by G. Schmid, J. O'Reilly, and K. Schomann. London: Edward Elgar.

Heckman, J. J. and J. A. Smith. (1997) "The sensitivity of experimental impact estimates: evidence from the National JTPA study." *NBER Working Paper* (6105).

Heckman, J. J. and J. A. Smith. (1998) "Evaluating the Welfare State." *NBER Working Paper* (6542).

Heckman, J. J. and J. A. Smith. (1999) "The pre-programme earnings dip and the determinants of participation in a social programme. Implications for simple programme evaluation strategies." *The Economic Journal, 109*, 313-348.

Heckman, J. J., J. A. Smith, and N. Clements. (1997) "Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts." *Review of Economic Studies, 64*, 487-535.

Heckman, J. J., J. A. Smith, and C. Taber. (1998) "Accounting for dropouts in evaluations of social programs." *The Review of Economics and Statistics, 80*, 1-14.

Hedges, L. V. (1997) "The promise of replication in labour economics." *Labour Economics, 4*, 111-114.

Hedges, L. V., R. Laine, D, and R. Greenwald. (1994) "Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes." *Educational Researcher, 23*, 5-14.

Hillage, J. and E. Pollard. (1998) "Employability: developing a framework for policy analysis." DfEE Research Report, 85.

Hunter, J. E. and F. L. Schmidt. (1990) *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Newbury Park: Sage Publications.

Hutchinson, G. and A. Church. (1989) "Wages, unions, the Youth Training Scheme and the Young Workers Scheme." *Scottish Journal of Political Economy, 36*, 160-182.

Intriligator, M. D., R. G. Bodkin, and C. Hsiao. (1989) *Econometric Models, Techniques and Applications*: Prentice-Hall.

Janoski, T. (1996) "Explaining State Intervention to Prevent Unemployment: The impact of Institutions on Active labour Market Policy Expenditure in 18 Countries." pp. 37-88 in *International Handbook of Labour Market Policy and Evaluation*, edited by G. Schmid, J. O'Reilly, and K. Schomann. London: Edward Elgar.

Jarrell, S. B. and T. D. Stanley. (1990) "A meta-analysis of the union-nonunion wage gap." *Industrial and Labor Relations Review, 44*, 54-67.

Jones, L. and D. Fiske. (1953) "Models for testing the significance of combined results." *Psychological Bulletin, 50*, 375-382.

Jones, I. (1988) "An evaluation of YTS." *Oxford Review of Economic Policy, 4*, 54-71.

Junankar, P. N. (1987) *From School to Unemployment? The Labour Market for Young People*. London: MacMillan Press.

Keane, M. and R. Moffitt. (1998) "A structural model of multiple welfare program participation and labor supply." *International Economic Review, 39*, 553-589.

Keep, E. and K. Mayhew. (1999) "The assessment: knowledge, skills, and competitiveness." *Oxford Review of Economic Policy, 15*, 1-15.

Kendall, M. and A. Stuart. (1979) *The Advanced Theory of Statistics*, vol. 1. London: Charles Griffin and Co.

Kenny, L. W., L.-F. Lee, G. S. Maddala, and R. P. Trost. (1979) "Returns to college education: An investigation of self-selection bias based on the project talent data." *International Economic Review, 20*, 775-789.

King, M. (1980) "An econometric model of tenure choice and demand for housing as a joint decision." *Journal of Public Economics, 14*, 137-159.

Koenker, R. and G. Basset. (1978) "Regression Quantiles." *Econometrica, 46*, 33-50.

Korenman, S. and D. Neumach. (1997) "Cohort crowding and youth labour markets: A cross national analysis." *NBER Working Paper* (6031).

LaLonde, R. and R. Maynard. (1987) "How precise are evaluations of employment and training programs: evidence from a field experiment." *Evaluation Review, 11*, 428-451.

LaLonde, R. J. (1986) "Evaluating the econometric evaluations of training programs with experimental data." *The American Economic Review, 76*, 604-619.

LaLonde, R. J. (1995) "The promise of public sector-sponsored training programmes." *Journal of Economic Perspectives, 9*, 149-168.

Lancaster, T. (1990) *The Econometric Analysis of Transition Data.* Cambridge: Cambridge University Press

Leamer, E. E. (1989) "Sensitivity analyses would help." *American Economic Review, 75*, 308-313.

Lee, L.-F. (1978) "Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables." *International Economic Review, 19*, 415-433.

Lee, L.-F. (1982) "Some approaches to the correction of selectivity bias." *Review of Economic Studies, 49*, 355-372.

Lee, L.-F. (1983) "Generalised econometric models with selectivity." *Econometrica, 51*, 507-512.

Lee, L.-F. (1993) "Multivatiate Tobit Models in Econometrics." pp. 507-512 in *Handbook of Statistics*, vol. 11, edited by G. S. Maddala, C. R. Rao, and H. D. Vinod: Elsevier.

Leung, S. F. and S. Yu. (1996) "On the choice between sample selection and two-part models." *Journal of Econometrics, 72*, 197-229.

Lewis, H. G. (1974) "Comments on selectivity biases in wage comparisons." *Journal of Political Economy, 82*, 1145-1155.

Lindley, R. M. (1996) "The school-to-work transition in the United Kingdom." *International Labour Review, 135*, 159-180.

Little, R. J. A. (1985) "A note about models for selectivity bias." *Econometrica, 53*, 1469-1474.

Lynch, L. M. (1992) "Private sector training and the earnings of young workers." *American Economic Review, 82*, 299-312.

Lynch, L. M. (1993) "The economics of youth training in the United States." *The Economic Journal, 103*, 1292-1302.

Maddala, G. S. (1983) "Limited Dependent and Qualitative Variables in Econometrics." in *Econometric Society Monographs*, vol. 3. Boston: Cambridge University Press.

Maddala, G. S. (1986) "Disequilibrium, Self-Selection, and Switching Models." pp. 1633-1688 in *Handbook of Econometrics*, vol. 3, edited by Z. Griliches and M. D. Intriligator: Elsevier.

Maddala, G. S. (1995) "Specification tests in limited dependent variable models." pp. 3-49 in *Advances in Econometrics and Quantitative Economics*, edited by G. S. Maddala, P. C. B. Phillips, and T. N. Srinivasan. Oxford: Blackwell.

Main, B. G. M. (1985) "School-leaver unemployment and the Youth Opportunities Programme in Scotland." *Oxford Economic Papers, 37*, 426-447.

Main, B. G. M. (1987) "The wage expectations and unemployment experience of school leavers." *Scottish Journal of Political Economy, 34*, 349-367.

Main, B. G. M. (1988) "The reporting of gross and net earnings in a postal survey: evidence from the Scottish School Leavers Survey." *Quality and Quantity, 22*, 99-110.

Main, B. G. M. (1991) "The effect of the Youth Training Scheme on employment probability." *Applied Economics, 23*, 367-372.

Main, B. G. M. and D. Raffe. (1983) "The 'Transition from School to Work' in 1980/81: A dynamic account." *British Educational Research Journal, 9*, 57-70.

Main, B. G. M. and D. Raffe. (1983) "Determinants of employment and unemployment among school leavers. Evidence from the 1979 Survey of Scottish School Leavers." *Scottish Journal of Political Economy, 30*, 1-17.

Main, B. G. M. and M. A. Shelly. (1988) "School leavers and the search for employment." *Oxford Economic Papers, 40*, 487-504.

Main, B. G. M. and M. A. Shelly. (1990) "The effectiveness of the Youth Training Scheme as a manpower policy." *Economica, 57*, 495-514.

Manski, C. F. (1990) "Anatomy of the selection problem." *The Journal of Human Resources, 24*, 343-360.

Manski, C. F. and I. Garfinkel. (1992) *Evaluating Welfare and Training Programs.* Cambridge, Massachusetts: Harvard University Press.

Manski, C. F. and S. R. Lerman. (1977) "The estimation of choice probabilities from choice based samples." *Econometrica, 45*, 1977-1988.

Marsden, D. W. and P. Ryan. (1990) "Institutional aspects of youth unemployment and training policy in Britain." *British Journal of Industrial Relations, 28*, 351-370.

Marcotte, D. E. (1999) "Learning in the labor market: The changing importance of education and training after 'formal' schooling ends." *NCRVE* (MDS-1275).

Martin, J. P. (1998) "What works among active labour market polices: Evidence from OECD countries' experiences." *Labour Market and Social Policy Occasional Papers* (No. 35, OECD, Paris).

Maynard, R. A. (1993) "Evaluating employment and training programmes: Lessons from the USA." *International Journal of Manpower, 14 (2/3)*, 94-104.

McNabb, R. and S. Richardson. (1989) "Earnings, education and experience: Is Australia different?" *Australian Economic Papers*: 57-75.

McNabb, R. and K. Whitfield. (1994) *The Market for Training*. Aldershot: Avebury.

Mealli, F., S. Pudney, and J. Thomas. (1995) "Youth employment and the optimal structure of youth training: an econometric analysis." *Leicester University Discussion Papers in Economics,* Leicester.

Mealli, F., S. Pudney, and J. Thomas. (1996) "Training duration and post-training outcomes: A duration-limited competing risks model." *The Economic Journal, 106*, 422-433.

Melino, A. (1982) "Testing for sample selection bias." *Review of Economic Studies, 49*, 151-153.

Meyer, B. D. (1995) "Natural and quasi-experiments in economics." *Journal of Business and Economic Statistics* (April) 151-161.

Micklewright, J. (1989) "Choice at Sixteen." *Economica, 56*, 25-39.

Miller, P. W. (1989) "The structure of Aboriginal and Non-Aboriginal youth unemployment." *Australian Economic Papers*: 39-56.

Miller, P. W. (1990) "Training in the youth labour market." *Labour Economics and Productivity, 2*, 1-26.

Miller, P. W. (1995) "The Australian Longitudinal Survey and Australian Youth Survey." *The Australian Economic Review*: 123-129.

Miller, P. W. and P. Volker. (1987) "The youth labour market in Australia." *The Economic Record* (September), 203-219.

Mincer, J. (1974) *Schooling, Experience and Earnings*. New York: Columbia University Press.

Mizen, P. (1995) *The State, Young People and Youth Training: In and Against the Training State*. London: Mansell.

Moffitt, R. (1987) "Symposium on the econometric evaluation of manpower training programs: introduction." *Journal of Human Resources, 22*, 149-156.

Moffitt, R. (1991) "Program evaluation with non-experimental data." *Evaluation Review, 15*, 149-156.

Mroz, T. A. (1987) "The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions." *Econometrica, 55*, 765-799.

Murphy, P. D., P. J. Sloane, and D. H. Blackaby. (1992) "The effects of trade unions on the distribution of earnings: A sample selectivity approach." *Oxford Bulletin of Statistics, 54*, 517-542.

Nathan, R. (1991) "Evaluation Strategies with Particular Emphasis on Demonstration Studies." pp. 21-42 in *Evaluating Labour Market and Social Programmes: The State of a Complex Art*. Paris: OECD.

National Audit Office [NAO] (2000) "Scottish Enterprise: Skillseekers for Young People." The Stationery Office, Edinburgh.

Nawata, K. (1994) "Estimation of sample selection bias models by the maximum likelihood estimator and Heckman's two-step estimator." *Economic Letters, 45*, 33-40.

Nelson, F. D. (1984) "Efficiency of the two-step estimator for models with endogenous sample selection." *Journal of Econometrics, 24*, 181-196.

Neumark, D. and W. Wascher. (1994) "Minimum wage effects on employment and school enrolment." *NBER Working Paper* (4679).

Neumark, D. and W. Wascher. (1998) "Minimum wage and training revisited." *NBER Working Paper* (6651).

Newey, W. K., J. L. Powell, and J. R. Walker. (1990) "Semiparametric estimation of selection models: some empirical results." *American Economic Review, AEA Papers and Proceedings, 80*, 324-328.

O'Connell, P. J. and F. McGinnity. (1997) "What works, who works? The employment and earnings effects of active labour market programmes among young people in Ireland." *Work, Employment and Society, 11*, 639-661.

Office for National Statistics [ONS]. (various) "Labour Market Trends." HMSO, London.

O'Higgins, N. (1994) "YTS, employment, and sample selection bias." *Oxford Economic Papers, 46*, 605-628.

O'Higgins, N. (1997) "The challenge of youth unemployment." *ILO Employment and Training Papers (7)* Geneva

Oi, W. (1962) "Labor as a quasi-fixed factor." *Journal of Political Economy, 70*, 538-555.

Olsen, R. J. (1980) "A least squares correction for selectivity bias." *Econometrica, 48*, 1815-1821.

Olsen, R. J. (1982) "Distributional tests for selectivity bias and a more robust likelihood estimator." *International Economic Review, 23*, 223-241.

Organisation for Economic Cooperation and Development [OECD]. (1990) *Labour Market Policies for the 1990s*, vol. 46. Paris: OECD.

OECD. (1991) *Evaluating Labour Market and Social Programmes: The State of a Complex Art*. Paris: OECD.

OECD. (1998) *Education at a Glance: OECD Indicators*. Paris: Centre for Educational Research and Innovation, OECD.

OECD. (1998) *Labour Force Statistics, 1977-1997*, vol. 46. Paris: OECD.

OECD. (1999) *Preparing Youth for the 21st Century: The Transition from Education to the Labour Market*. Paris: OECD.

OECD. (various) *Economic Outlook*, vol. 46. Paris: OECD.

Orme, C. (1995) "On the use of certain artificial regressions in certain micro-econometric models." *Econometric Theory, 11*, 290-305.

Orr, L. L. (1999) *Social Experiments: Evaluating Public Programs with Experimental methods*, vol. 46. Thousand Oaks: Sage.

Paarsch, H. J. (1984) "A Monte Carlo comparison of estimators for censored regression models." *Journal of Econometrics, 24*, 197-213.

Patterson, L. and D. Raffe. (1995) "'Staying-on' in full-time education in Scotland, 1985-1991." *Oxford Review of Education, 21*, 3-23.

Payne, J. (1987) "Does unemployment run in families? Some findings from the General Household Survey." *Sociology, 2*, 199-214.

Payne, J. (1995) "Options at 16 and outcomes at 24: A comparison of academic and vocational education and training routes." Policy Studies Institute, D*f*EE Youth Cohort Series, 35.

Payne, J. (1999) "Recent Changes in School-to-Work Transition in England and Wales." pp. 475-501 in *International Perspectives on the School-to-Work Transition*, edited by D. Stern and D. A. Wagner. Cresskill: Hampton Press.

Payne, J., B. Casey, C. Payne, and S. Connolly. (1996) *Long-Term Unemployment: Individual Risks and Outcomes*, vol. 2. London: Policy Studies Institute.

Pearson, E. S. (1938) "The probability integral transformation for testing goodness of fit and combining tests of significance." *Biometrika, 30*, 134-148.

Pearson, K. (1933) "On a method of determining whether a sample of size *n* supposed to have been drawn from a parent population having a known probability integral has probably been drawn at random." *Biometrika, 25*, 379-410.

Pierre, G. (1999) "A framework for active labour market policy evaluation." *ILO Employment and Training Papers (49)*, Geneva.

Powell, J. L. (1986) "Censored regression quantiles." *Journal of Econometrics, 32*, 143-155.

Puhani, P. A. (2000) "The Heckman correction for sample selection and its critique." *Journal of Economic Surveys, 14*, 53-68.

Raffe, D. (1990) "The Context of the Youth Training Scheme: An analysis of its strategy and development." pp. 58-75 in *Training and its Alternatives*.

Raffe, D., A. Biggart, J. Fairgrieve, C. Howieson, J. Rodger, and S. Burnisten. (1998) "Thematic Review of the Transition from Initial Education to Working Life: United Kingdom, background report." OECD, Paris.

Ray, D., W. Beswick, C. Lawson, C. O'Brien, and S. Madigan. (2000) "Attrition in apprenticeships." DETYA, REB 1/100, Canberra.

Richardson, J. (1997) "Can active labour market policy work? Some theoretical considerations." *Centre for Economic Performance* (Discussion Paper No. 331).

Riddell, C. (1991) "Evaluation of Manpower and Training Programmes." pp. 43-72 in *Evaluating Labour Market and Social Programmes: The State of a Complex Art*. Paris: OECD.

Robinson, P. (1996) "Labour Market Studies: United Kingdom." Report for Directorate-General for Employment, Industrial Relations and Social Affairs, European Commission.

Robson, J. (1993) "The Youth Cohort Study: methodological introduction." *Education Gazette*: 251-258.

Rosenbaum, P. R. (1984) "The consequences of adjustment for a concomitant variable that has been affected by the treatment." *Journal of Royal Statistical Society: A, 147*, 656-666.

Rosenbaum, P. R. and D. B. Rubin. (1983) "Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome." *Journal of Royal Statistical Society: B, 45*, 212-218.

Rosenthal, R. (1984) "Meta-Analysis Procedures for Social Research." in *Applied Social Research Methods Series*, vol. 6, edited by L. Bickman. Beverley Hills: Sage Publications.

Rosenthal, R. (1997) "Some issues in the replication of social science research." *Labour Economics, 4*, 121-123.

Roy, A. D. (1951) "Some thoughts on the distribution of earnings." *Oxford Economic Papers, 3*, 135-146.

Ryan, P. (1995) "Trade union policies towards the Youth Training Scheme: Patterns and causes." *British Journal of Industrial Relations, 33*, 1-33.

Ryan, P. and C. F. Buchtemann. (1996) "The school-to-work transition." pp. 308-347 in *International Handbook of Labour Market Policy and Evaluation*, edited by G. Schmid, J. O'Reilly, and K. Schomann. London: Edward Elgar.

Sako, M. and R. Dore. (1986) "How the Youth Training Scheme helps employers." *Employment Gazette, 94*, 195-204.

Schmertmann, C. P. (1994) "Selectivity bias correction methods in polychotomous sample selection models." *Journal of Econometrics, 60*, 101-132.

Schofeild, K. (1999a) "Independent Investigation into the Quality of Training in Queensland's Traineeship System." Department of Employment, Training and Industrial Relations, Brisbane.

Schofeild, K. (1999b) "A Risky Business: Review of the Quality of Tasmania's Traineeship System." Office of Vocational Education and Training, Hobart.

Schofeild, K. (2000) "Report of the Independent Review of the Quality of Training in Victoria's Apprenticeship and Traineeship System." Office of Post Compulsory Education, Training and Employment, Melbourne.

Schomann, K. (1996) "Longitudinal Designs in Evaluation Studies." pp. 115-142 in *International Handbook of Labour Market Policy and Evaluation*, edited by G. Schmid, J. O'Reilly, and K. Schomann, Cheltenham.

Sen, A. (1997) "Inequality, unemployment and contemporary Europe." *International Labour Review, 136*, 155-171.

Smith, A. (1976 Reprint) *An Inquiry into the Nature and Causes of The Wealth of Nations*. Chicago: The University of Chicago Press.

Smith, K. V. and J.-C. Huang. (1995) "Can markets value air quality? A meta-analysis of hedonic property value models." *Journal of Political Economy, 103*, 209-227.

Stanley, T. D. and S. B. Jarrell. (1989) "Meta-regression analysis: A quantitative method for literature surveys." *Journal of Economic Surveys, 3*, 161-170.

Steedman, H. (1993) "The economics of youth training in Germany." *The Economic Journal, 103*, 1279-1291.

Steedman, H. and J. Hawkins. (1994) "Shifting foundations: The impact of NVQs on youth training for the building trades." *National Institute Economic Review*: 93-103.

Stern, D., T. Bailey, and D. Merritt. (1996) "School-to-work policy insights from recent international devleopments." *NCRVE* (MDS-950).

Stern, D. and D. A. Wagner. (1999) "Introduction: School-to-Work Policies in Industrialized Countries as Responses to Push and Pull." pp. 69-124 in *International Perspectives on the School-to-Work Transition*, edited by D. Stern and D. A. Wagner. Cresskill: Hampton Press.

Stevens, J. and R. Mackay. (1991) "Training and Competitiveness." in *Policy Issues Series*. London: National Economic Development Office.

Stevens, J. and T. Walsh. (1991) "Training and Competitiveness." in *Training and Competitiveness*, edited by J. Stevens and R. Mackay. London: National Economic Development Office.

Stevens, M. (1999) "Human capital theory and UK vocational training policy." *Oxford Review of Economic Policy, 15*, 16-32.

Stolzenberg, R. M. and D. A. Relles. (1997) "Tools for intuition about sample selection bias and its correction." *American Sociological Review, 62*, 494-507.

Stromsdorfer, E. W. (1987) "Economic evaluation of the Comprehensive Employment and Training Act: an overview of recent findings and advances in evaluation methods." *Education Review, 11*, 387-394.

Strube, M. J. and R. H. Miller. (1986) "Comparison of power rates for combined probability procedures: A simulation study." *Psychological Bulletin, 99*, 407-415.

Sweet, R. (1999) "An Overview of School-to-Work Policy Arrangements in Australia." pp. 69-124 in *International Perspectives on the School-to-Work Transition*, edited by D. Stern and D. A. Wagner. Cresskill: Hampton Press.

Tan, H., D. Chapman, C. Peterson, and A. Booth. (1991) "Youth training in the United States, Britain, and Australia." RAND, Santa Monica.

The Scottish Executive. (1999) *Leaver Destinations from Scottish Secondary Schools 1996/97 to 1998/99.*

The Scottish Office. (1998) *Opportunity Scotland: A Paper on Lifelong Learning.* London: The Stationery Office, Cm 4048, HMSO.

The Scottish Office. (1998) *Opportunities and Choices.* London: The Stationery Office.

The Scottish Office. (various) *Leaver Destinations from Scottish Secondary Schools.*

Thurow, L. (1975) *Generating Inequality: The Distributional Mechanisms of the Economy,* New York: Basic Books).

Thurow, L. (1994) "New game, new rules, new strategies." *RSA Journal* (CXLII), 50-56.

Trochim, W. M. K. (1999) *Research Methods Knowledge Base.* Ithaca, New York: Cornell University.

Vella, F. (1992) "Simple tests for sample selection bias in censored and discrete choice models." *Journal of Applied Econometrics, 7*, 413-421.

Vella, F. (1997) "Estimating models with sample selection bias: A survey." *The Journal of Human Resources, 33*, 127-169.

Vella, F. and T. Karmel. (1999) "Evaluating the impact of educational expansion on the occupational status of youth." *Australian Economic Papers* (September), 310-327.

Warburton, W. P. (1996) "What went wrong in the CETA evaluations?" *Canadian Journal of Economics, 29*, S105-108.

White, M. and J. Lakey. (1992) *The Restart Effect: Does Active Labour Market Policy Reduce Unemployment?* vol. 18. London: Policy Studies Institute.

White, H. (1980) "A heteroscedasticity consistent covariance matrix estimator and a direct test of heteroscedasticity. " *Econometrica, 48*, 817-818.

Whitfield, K. and C. Bourlakis. (1991) "An empirical analysis of YTS, employment and earnings." *Journal of Economic Studies, 18*, 42-56.

Willis, R. J. and S. Rosen. (1980) "Education and Self-Selection." pp. 413-442 in *Evaluation Studies Review Annual*, vol. 5, edited by E. W. Stromsdorfer and G. Farkas. Beverley Hills: Sage.

Winer, B. J. (1971) *Statistical Principles in Experimental Design*, vol. 18. New York: McGraw-Hill.

Winship, C. and R. D. Mare. (1992) "Models for sample selection bias." *Annual Review of Sociology, 18*, 327-350.

Wolf, F. M. (1986) *Meta-Analysis: Quantitative Methods for Research Synthesis.* Beverley Hills: Sage Publications.

Wooden, M. (1995) "Training data and statistics in Australia." *The Australian Economic Review*: 116-120.

Yates, F. and W. G. Cochrane. (1938) "The analysis of groups of experiments." *Journal of Agricultural Science, 28*, 556-580.

Electronic Sources

Australia: Government and Statistics

http://www.abs.gov.au/
http://www.anta.gov.au/
http://www.detya.gov.au/
http://www.fed.gov.au/
http://ssda.anu.edu.au/

United Kingdom: Government and Statistics

http://www.dfee.gov.uk/
http://www.scotland.gov.uk/
http://www.statistics.gov.uk/
http://www.ukonline.gov.uk/

European Commission:

http://europa.eu.int/

International Organisations:

http://www.ilo.org/
http://www.oecd.org/

Appendix

Table A1: Key Youth/Adult Labour Market Statistics in Selected OECD Countries, 1988 and 1998 (%)

		Unemployment Rate		*Labour Force Participation Rate*		*Employment/Population Ratio* [1]	
		1	*2*	*3*	*4*	*5*	*6*
Country \ *Age Group*		15-24	25-54	15-24	25-54	15-24	25-54
Australia	1988	12.8	4.9	69.1	77.8	61.1	75.8
	1998	14.5	6.3	67.6	80.0	57.8	75.0
Canada	1988	11.9	6.9	70.1	83.7	60.5	78.4
	1998	15.2	7.1	62.0	84.3	52.6	78.3
Finland	1988	7.8	3.8	58.1	90.1	54.4	86.9
	1998	22.0	9.5	49.7	87.1	38.8	78.9
France	1988	21.6	8.4	39.0	83.4	29.5	77.4
	1998	25.4	10.8	28.0	86.2	20.9	76.8
Germany [b]	1988	7.7	7.1	61.5	77.8	56.4	73.6
	1998	9.4	7.7	49.6	83.5	45.0	77.0
Ireland	1988	23.9	14.9	52.5	67.4	41.5	60.2
	1998	11.5	7.3	48.6	76.4	43.0	70.9
Italy	1988	34.5	7.4	45.1	69.2	29.8	64.9
	1998	32.1	9.6	37.5	72.8	25.4	65.9
Japan	1988	4.9	2.0	42.6	79.6	42.2	79.6
	1998	7.7	3.4	48.3	82.1	44.6	79.2
Netherlands	1988	13.6	8.1	59.8	74.9	54.5	66.5
	1998	8.2	3.6	68.0	82.3	62.5	79.3
New Zealand	1988	10.8	4.1	69.0	81.4	58.3	76.3
	1998	14.6	6.1	65.2	81.8	55.7	76.8
Norway [a]	1988	8.2	2.1	66.9	86.9	54.7	82.5
	1998	9.5	2.3	63.8	87.8	57.7	85.7
Portugal	1988	12.9	4.4	62.2	79.9	51.2	77.4
	1998	9.5	4.1	47.4	83.8	42.9	80.3
Spain [a]	1988	37.1	14.7	58.3	68.1	34.7	61.0
	1998	34.1	16.5	46.4	75.6	30.6	63.1
Sweden [a]	1988	4.2	1.4	68.3	92.4	64.5	91.2
	1998	16.8	7.6	50.0	88.0	41.6	81.3
UK [a]	1988	12.8	7.5	77.7	82.7	70.1	79.0
	1998	12.3	5.0	69.5	83.3	61.0	79.1
US [a]	1988	11.0	4.5	68.4	82.9	59.8	79.8
	1998	10.4	3.5	65.9	84.1	59.0	81.1
OECD Avg. [c]	1988	14.4	6.0	57.4	78.5	49.4	75.3
	1998	12.8	5.7	51.7	80.1	45.1	75.5

Source: 1988 figures – OECD Labour Force Statistics, 1977-1997
1998 figures – OECD Employment Outlook, 1999 June

[a] *Youth measures are for 16 to 24 years of age. All the measures for the United Kingdom exclude Northern Ireland.*

[b] *1988 figures are for west Germany only.*

[c] *Unweighted average includes all 16 countries listed plus Austria, Belgium, Czech Republic, Denmark, Greece, Hungary, Iceland, Luxembourg, Mexico, Poland, Switzerland and Turkey.*

[1] *1988 Employment/population ratio figures refer to 1989*

Table A2: Public expenditure on labour market policy as a percentage of GDP in selected OECD countries, 1985 - 2000

Measure* / Country	Training: *PES & Admin*	Training: *Adult Measures*	Training: ***Youth Measures***	Job Subsidies	Disabled	**Total Active Measures**	U-benefit	Early Retirement	**Total Passive Measures**	Total Spending	**Youth** U-rate [a]	Aggregate U-rate
	1	*2*	*3*	*4*	*5*	*(1-5)*	*6*	*7*	*(6-7)*		%	%
Australia												
1985/86	0.11	0.02	**0.07**	0.19	..	**0.39**	1.31	..	**1.31**	1.70	**14.5**	8.0
1990/91	0.07	0.06	**0.04**	0.04	0.04	**0.25**	1.07	..	**1.07**	1.32	**13.3**	6.9
1995/96	0.24	0.15	**0.06**	0.31	0.07	**0.83**	1.28	..	**1.28**	2.10	**14.4**	8.5
1999/2000	0.20	0.02	**0.07**	0.11	0.05	**0.46**	1.05	..	**1.05**	1.51	**12.3**	6.6
Canada												
1985/86	0.24	0.34	**0.03**	0.02	..	**0.63**	1.87	..	**1.87**	2.50	**16.3**	10.4
1990/91	0.21	0.27	**0.02**	0.02	..	**0.52**	1.91	..	**1.91**	2.43	**12.7**	8.2
1995/96	0.21	0.25	**0.02**	0.06	0.02	**0.56**	1.28	0.01	**1.29**	1.85	**15.6**	9.5
1999/2000	0.20	0.17	**0.03**	0.08	0.03	**0.50**	0.98	..	**0.98**	1.49	**12.6**	7.0
Finland												
1985	0.08	0.27	**0.05**	0.42	0.09	**0.91**	0.87	0.46	**1.34**	2.25	**9.1**	5.0
1990	0.10	0.25	**0.05**	0.42	0.12	**0.94**	0.62	0.49	**1.11**	2.05	**8.9**	3.4
1995	0.16	0.45	**0.16**	0.68	0.14	**1.58**	3.57	0.44	**4.01**	5.59	**26.9**	15.3
2000	0.12	0.35	**0.19**	0.32	0.09	**1.08**	1.75	0.47	**2.22**	3.30	**21.6**	9.2
France												
1985	0.13	0.25	**0.17**	0.06	0.05	**0.67**	1.20	1.21	**2.41**	3.07	**25.6**	10.2
1990	0.13	0.33	**0.22**	0.07	0.06	**0.81**	1.32	0.56	**1.88**	2.68	**19.1**	9.0
1995	0.15	0.38	**0.27**	0.40	0.10	**1.31**	1.43	0.36	**1.79**	3.09	**25.9**	11.7
1999/2000	0.17	0.28	**0.41**	0.41	0.09	**1.36**	1.47	0.29	**1.76**	3.12	**20.7**	8.5
Germany												
1985	0.21	0.20	**0.05**	0.17	0.19	**0.81**	1.41	0.01	**1.42**	2.23	**9.9**	7.2
1990	0.22	0.37	**0.04**	0.15	0.23	**1.01**	1.14	0.02	**1.16**	2.17	**4.6**	4.8
1995	0.23	0.38	**0.06**	0.44	0.26	**1.37**	2.09	0.29	**2.38**	3.75	**8.5**	8.2
2000	0.23	0.34	**0.08**	0.31	0.27	**1.23**	1.88	0.01	**1.89**	3.13	**7.7**	7.7

Table A2 continued:

Measure* / Country	Training: PES & Admin	Training: Adult Measures	Training: *Youth Measures*	Job Subsidies	Disabled	**Total Active Measures**	U-benefit	Early Retirement	**Total Passive Measures**	Total Spending	**Youth** U-rate [a]	Aggregate U-rate
	1	*2*	*3*	*4*	*5*	*(1-5)*	*6*	*7*	*(6-7)*		*%*	*%*
Ireland												
1985	0.14	0.66	**0.51**	0.19	..	**1.51**	3.69	0.00	**3.69**	5.20	**24.1**	17.0
1990	0.16	0.49	**0.39**	0.29	0.14	**1.47**	2.79	0.05	**2.84**	4.31	**17.6**	13.4
1995	0.27	0.22	**0.27**	0.86	0.09	**1.68**	2.57	0.14	**2.71**	4.39	**19.1**	12.3
Italy												
1985	0.08	0.06	**0.32**	..	..	**0.45**	0.75	0.28	**1.04**	1.49	**33.9**	9.6
1990	0.08	0.03	**0.69**	..	..	**0.80**	0.40	0.32	**0.72**	1.51	**28.9**	10.3
1995	0.04	0.01	**0.39**	0.69	..	**1.13**	0.68	0.19	**0.87**	2.00	**32.8**	12.3
1999	0.04	0.25	**0.27**	0.53	..	**1.10**	0.55	0.09	**0.64**	1.74	**31.5**	8.4
Japan												
1985/86	0.03	0.03	**..**	0.10	0.01	**0.17**	0.40	..	**0.40**	0.57	**5.2**	2.8
1990/91	0.02	0.03	**..**	0.07	0.01	**0.13**	0.32	..	**0.32**	0.45	**4.3**	2.1
1995/96	0.03	0.03	**..**	0.06	..	**0.13**	0.39	..	**0.39**	0.51	**6.1**	3.2
1998/1999	0.03	0.03	**..**	0.02	..	**0.09**	0.52	..	**0.52**	0.61	**9.2**	5.1
Netherlands												
1985	0.08	0.19	**0.04**	0.06	0.72	**1.09**	3.24	..	**3.24**	4.33	**22.9**	10.6
1990	0.08	0.19	**0.06**	0.05	0.62	**1.04**	2.30	..	**2.30**	3.34	**11.1**	7.5
1995	0.37	0.40	**0.09**	0.10	0.55	**1.51**	3.21	..	**3.21**	4.72	**12.8**	6.9
2000	0.25	0.30	**0.04**	0.41	0.57	**1.57**	2.08	..	**2.08**	3.65	**5.3**	2.2
New Zealand												
1985/86	0.05	0.09	**0.01**	0.62	0.01	**0.79**	0.65	..	**0.65**	1.44	**7.8**	4.1
1990/91	0.16	0.39	**0.09**	0.19	0.06	**0.89**	1.90	..	**1.90**	2.79	**14.1**	7.9
1995/96	0.13	0.32	**0.09**	0.14	0.03	**0.71**	1.14	..	**1.14**	1.85	**11.9**	6.3
1999/2000	0.07	0.18	**0.14**	0.11	0.05	**0.55**	1.62	..	**1.62**	2.17	**13.2**	6.2

Table A2 continued:

Measure* / Country	Training			Job Subsidies	Disabled	**Total Active Measures**	U-benefit	Early Retirement	**Total Passive Measures**	Total Spending	**Youth U-rate** [a]	Aggregate U-rate
	PES & Admin	*Adult Measures*	***Youth Measures***									
	1	*2*	*3*	*4*	*5*	*(1-5)*	*6*	*7*	*(6-7)*		*%*	*%*
Norway [b]												
1985	0.12	0.11	**0.05**	0.21	0.18	**0.66**	0.50	..	**0.50**	1.17	**6.5**	2.6
1990	0.14	0.36	**0.12**	0.18	0.20	**1.00**	1.17	..	**1.17**	2.17	**11.8**	5.2
1995	0.18	0.23	**0.08**	0.22	0.63	**1.34**	1.10	..	**1.10**	2.44	**11.9**	5.0
2000	0.12	0.08	**0.01**	0.01	0.55	**0.77**	0.39	..	**0.39**	1.16	**10.2**	3.6
Portugal [c]												
1986	0.08	0.21	**0.04**	0.04	0.04	**0.41**	0.41	..	**0.41**	0.81	**19.3**	8.5
1990	0.11	0.14	**0.12**	0.09	0.07	**0.52**	0.32	..	**0.32**	0.84	**10.4**	4.6
1995	0.10	0.23	**0.34**	0.08	0.05	**0.68**	0.84	0.09	**0.93**	1.73	**16.0**	7.4
1998	0.11	0.30	**0.18**	0.09	0.01	**0.69**	0.67	0.16	**0.83**	1.52	**8.4**	3.2
Spain [b]												
1985	0.09	0.02	**0.00**	0.22	0.01	**0.34**	2.87	0.02	**2.89**	3.23	**42.8**	21.4
1990	0.13	0.10	**0.07**	0.45	0.00	**0.75**	2.42	..	**2.42**	3.17	**30.1**	15.9
1995	0.09	0.32	**0.09**	0.33	0.01	**0.84**	2.41	..	**2.41**	3.25	**40.3**	22.0
2000	0.05	0.29	**0.06**	0.41	0.03	**0.98**	1.34	..	**1.34**	2.32	**25.5**	9.7
Sweden [b]												
1985/86	0.25	0.50	**0.21**	0.43	0.72	**2.11**	0.75	0.12	**0.87**	2.97	**6.8**	2.8
1990/91	0.21	0.54	**0.07**	0.13	0.76	**1.71**	0.81	0.08	**0.89**	2.60	**4.5**	1.5
1995/96	0.26	0.55	**0.02**	0.82	0.70	**2.36**	2.26	..	**2.26**	4.62	**20.6**	8.8
2000	0.26	0.31	**0.02**	0.27	0.52	**1.38**	1.34	..	**1.34**	2.72	**11.9**	6.3
UK [b]												
1985/86	0.14	0.09	**0.25**	0.22	0.03	**0.72**	2.01	0.05	**2.06**	2.79	**17.9**	11.2
1990/91	0.14	0.22	**0.18**	0.02	0.03	**0.59**	0.90	..	**0.90**	1.49	**10.1**	6.8
1995/96	0.20	0.10	**0.12**	0.02	0.03	**0.46**	1.33	..	**1.33**	1.33	**15.3**	8.7
1999/2000	0.13	0.05	**0.15**	0.01	0.02	**0.37**	0.58	..	**0.58**	0.94	**11.8**	6.1

Table A2 continued:

*Measure** / *Country*	Training			Job Subsidies	Disabled	**Total Active Measures**	U-benefit	Early Retirement	**Total Passive Measures**	Total Spending	**Youth U-rate**[a]	Aggregate U-rate
	PES & Admin	*Adult Measures*	***Youth Measures***									
	1	*2*	*3*	*4*	*5*	*(1-5)*	*6*	*7*	*(6-7)*		*%*	*%*
US[b]												
1985/86	0.07	0.12	**0.04**	0.02	0.04	**0.30**	0.61	..	**0.61**	0.91	**13.3**	7.1
1990/91	0.08	0.09	**0.03**	0.01	0.04	**0.25**	0.60	..	**0.60**	0.85	**11.1**	5.4
1995/96	0.08	0.04	**0.03**	0.01	0.04	**0.19**	0.34	..	**0.34**	0.53	**12.1**	5.6
1999/2000	0.04	0.04	**0.03**	0.01	0.03	**0.15**	0.23	..	**0.23**	0.38	**9.3**	3.9

Notes to table:

Source: OECD Employment Outlook, 1992-1998
[a] *15 to 24 years of age*
[b] *Youth unemployment rate is for 16 to 24 years of age. All the measures for the United Kingdom exclude Northern Ireland.*
[c] *1986 was the first year data was available for this country, youth unemployment rate is for 14 to 24 years of age.*
.. Nil or less tan ½ of last digit used
Notes on measures:

1. *Public employment services and administration (PES & Admin) includes: placement, counselling and vocational guidance; job-search courses and related counselling; support of geographic mobility to aid job search and all administration costs including benefit agencies and job centres.*
2. *Labour market training (Adult Measures) includes: training for unemployed workers and workers at risk for reasons of labour market policy.*
3. *Youth measures (Youth Measures), includes measures for unemployed and disadvantaged youth, support of apprenticeship and related forms of general youth training.*
4. *Direct job creation and employment subsidies (Job Subsidies), subsidies targeted to the unemployed. Grants to enterprises for capital costs are not included.*
5. *Measures for the disabled (Disabled), includes vocational rehabilitation and work for the disabled*
6. *Unemployment compensation (U-benefit)*
7. *Early retirement for labour market reasons (Early Retirement)*

Table A3: Description of SYPS Variables

Personal background
Birth date
Income
Welfare

School Experience
Detail on:
- Attitude to lessons
- Resources and support
- Company kept
- Study habits
- Subjects studied
- Examinations attempted
- Examinations obtained
- Type of school
- Living arrangements
- Work experience
- Part-time job at school
- If bullied
- If punished
- If played truant
- Reasons for staying on
- Reasons for leaving
- Date of leaving
- Career guidance
- Any return details

Post-school experience
Where living
Living arrangements
Any children
Support
Employment status
Occupational classification
Industrial classification
Occupational code
Number of jobs
Intentions
Interests

Detail on:
- Courses applied for
- Course taken/taking
- Attitude to courses
- Qualifications
- Apprenticeship
- Other training
- First job
- Unemployment

Family background
Parents:
- Current main activity
- Age when left school
- If self-employed
- Employment status
- Occupational
- Industrial classification
- Occupational code
- Age
- Education
- Home ownership

Number of siblings

Youth Training
- Number of courses
- First course
- Most recent course
- Duration & qualifications
- Remaining with employer
- Wage paid over allowance

In addition there are a number of derived variables based upon responses and constructed using other resources such as the SIC and SOC industrial and occupational codes.

Breakdown of the integer index of exam success TOTSCEP:

1	no awards at O/S or S at 6 - 7
2	0/S at 4 - 5 only
3	one 0/S at 1 - 3
4	two 0/S at 1 - 3
5	three 0/S at 1 - 3
6	four 0/S at 1 - 3
7	five 0/S at 1 - 3
8	six plus 0/S at 1 - 3
9	one higher pass
10	two higher passes
11	three higher passes
12	four higher passes
13	five higher passes
14	six plus higher passes

Source: CES, SYPS technical documentation

Table A4: Construction of SYPS sample

Selection Criteria	*Number of observations dropped*
Total number of valid cases in 91cohort	4450
1. Did not complete all follow-up sweeps	1693
2. Still in full-time higher education	1195
3. Still in other full-time education	85
4. Still in YTS or similar scheme	78
5. Looking after family/pregnant	71
6. Doing unpaid work (voluntary)	8
7. Unable to work (ill/disabled)	3
Indicated not in labour force due to:	
8. Doing part-time higher education	9
9. Labour force status inadequately described	15
Remaining sample:	**1293**

Notes:

As detailed in the text the data are from the fourth cohort of the SYPS administered by the Centre for Educational Sociology (CES) at Edinburgh University. The target population was 10% of all those that reached the minimum school leaving age (16 years) during the 1989/90 school year and collection was based on a quasi-random technique.

The Centre keeps the data as SPSS files(.sav). The majority of the initial manipulations in the text were thus carried out in SPSS before saving the data as Lotus 1-2-3 worksheets (.wks). Once saved in this format the is then able to be read directly into LIMDEP 7.0. Note that the 'write variable names' option needs to ticked.

Table A5: MLE: wage equations for school-leavers in employment, Autumn 1993

Descriptive variables	*YT*	*Non-YT*
Personal characteristics		
Female	-0.094 (2.34)	-0.168 (4.89)
Marital status	0.094 (1.38)	-0.003 (0.05)
School experience		
Unqualified	-0.044 (0.58)	0.097 (1.28)
1-3 O/S-Grades	0.015 (0.28)	0.141 (2.44)
4+ O/S-Grades	0.047 (0.79)	0.102 (1.72)
Highers	0.107 (1.30)	0.100 (1.60)
S4 Leaver	0.166 (1.11)	0.110 (2.15)
S5 Leaver	0.126 (0.88)	0.061 (1.41)
Labour market experience		
Local unemployment rate	-0.002 (0.36)	-0.001 (0.09)
Unemployment duration (mths)	-0.007 (1.71)	-0.007 (2.57)
Hours worked	-0.012 (4.80)	-0.010 (6.92)
Does Training	0.077 (1.91)	0.026 (0.75)
Self Employed	0.021 (0.97)	0.406 (4.90)
Industrial and occupational descriptors		
Mining and utilities	-0.042 (0.00)	0.258 (2.16)
Manufacturing	0.088 (0.73)	0.091 (0.97)
Construction	0.031 (0.26)	-0.114 (1.13)
Services	-0.060 (0.53)	-0.027 (0.30)
No industry given	-0.543 (0.00)	-0.072 (0.43)
Unskilled	-0.058 (0.96)	-0.165 (4.02)
Skilled Manual	-0.108 (2.11)	-0.069 (1.58)
Professional/Intermediate occupation	0.163 (2.21)	-0.019 (0.32)
No occupation given	0.434 (3.04)	0.057 (0.56)
Youth Training		
YT One (< 12 months)	0.096 (1.99)	..
YT Three (> 24 months)	-0.024 (0.53)	..
YT Employer	-0.082 (2.00)	..
YT Qualification	0.073 (1.79)	..
Still on YT 1993	-0.037 (0.77)	..
Constant	6.348 (25.25)	6.339 (43.07)
Log-likelihood	-587.37	-706.65
Variance parameters (col.1 – col.2)		
Sigma (1) - Sigma (0)	0.284 (11.67)	0.373 (21.51)
Rho (1, 2) - Rho (0, 2)	-0.302 (0.84)	0.746 (9.715)
N	*317 (908)*	*591 (908)*

**coefficient (t-statistic)*

Table A6: OLS: wage equations for Christmas 1990 (S4) and Christmas 1991 (S5) school-leavers in employment, Autumn 1993

Descriptive variables	*S4 Leavers*	*S4 & S5 Leavers*
Personal characteristics		
Female	-0.213 (5.45)	-0.167 (6.04)
Marital status	0.017 (0.30)	0.036 (0.86)
School experience		
Unqualified	0.016 (0.33)	-0.024 (0.54)
1-3 O/S-Grades	0.061 (1.45)	0.053 (1.48)
4+ O/S-Grades	0.051 (1.07)	0.072 (1.82)
Highers	..	0.115 (2.53)
S4 Leaver	..	0.048 (1.65)
Labour market experience		
Local unemployment rate	-0.008 (1.78)	-0.003 (1.08)
Unemployment duration (mths)	-0.004 (1.05)	-0.008 (2.61)
Hours worked	-0.012 (5.46)	-0.011 (7.32)
Does Training	0.060 (1.66)	0.027 (1.00)
Self Employed	0.106 (0.95)	0.211 (2.13)
Industrial and occupational descriptors		
Mining and utilities	0.387 (2.39)	0.335 (2.68)
Manufacturing	0.131 (1.28)	0.048 (0.57)
Construction	0.005 (0.05)	-0.090 (1.04)
Services	-0.016 (0.16)	-0.062 (0.75)
No industry given	-0.064 (0.34)	-0.215 (1.39)
Unskilled	-0.124 (2.53)	-0.094 (2.66)
Skilled Manual	-0.144 (2.92)	-0.078 (2.16)
Professional/Intermediate occupation	0.098 (1.45)	0.081 (1.76)
No occupation given	0.227 (1.92)	0.210 (2.47)
Youth Training		
YT experience	-0.126 (2.02)	-0.119 (2.28)
YT One (< 12 months)	0.120 (2.15)	0.115 (2.44)
YT Three (> 24 months)	-0.083 (1.54)	-0.045 (0.96)
YT Employer	-0.039 (0.88)	-0.062 (1.64)
YT Qualification	0.092 (1.80)	0.079 (1.80)
Still on YT 1993	0.093 (1.31)	-0.003 (0.05)
Constant	6.606 (46.17)	6.530 (56.65)
F- statistic	4.75	6.85
Mean of dependent variable	5.9434	5.9632
Standard deviation of dependent variable	0.336	0.339
Standard error of regression	0.302	0.308
N	*398*	*741*

**coefficient (t-statistic)*

Table A7: Probit/OLS (Heckman Model): wage equations for Christmas 1999 (S4) and Christmas 1991 (S4 & S5) school-leavers in employment, Autumn 1993

	S4 Leavers		*S4 & S5 Leavers*	
Descriptive variables	*YT*	*Non-YT*	*YT*	*Non-YT*
Personal characteristics				
Female	-0.139 (2.80)	-0.310 (5.06)	-0.094 (2.38)	-0.209 (5.30)
Marital status	0.046 (0.63)	-0.014 (0.17)	0.092 (1.51)	-0.011 (0.18)
School experience				
Unqualified	0.062 (0.88)	0.011 (0.12)	-0.013 (0.23)	-0.036 (0.50)
1-3 O/S-Grades	0.097 (1.47)	0.064 (0.81)	0.034 (0.74)	0.090 (1.65)
4+ O/S-Grades	0.171 (1.54)	0.017 (0.16)	0.065 (1.18)	0.125 (2.13)
Highers	..	..	0.118 (1.62)	0.143 (2.37)
S4 Leaver	..	..	0.042 (0.95)	0.043 (1.16)
Labour market experience				
Local unemployment rate	-0.014 (1.64)	-0.009 (0.98)	-0.002 (0.34)	-0.011 (1.94)
Unemployment duration (mths)	-0.003 (0.59)	-0.001 (0.20)	-0.007 (1.71)	-0.007 (1.84)
Hours worked	-0.011 (3.67)	-0.013 (4.57)	-0.012 (4.70)	-0.010 (5.68)
Does Training	0.060 (1.31)	0.052 (0.97)	0.066 (1.73)	0.010 (0.26)
Self Employed	-0.018 (0.13)	0.199 (1.21)	0.021 (0.16)	0.401 (2.82)
Industrial and occupational				
Mining and utilities	-0.095 (0.31)	0.556 (2.81)	-0.048 (0.15)	0.368 (2.51)
Manufacturing	0.136 (0.99)	0.157 (1.07)	0.091 (0.76)	0.042 (0.38)
Construction	0.058 (0.42)	-0.029 (0.20)	0.031 (0.25)	-0.147 (1.28)
Services	-0.021 (0.16)	0.014 (0.10)	-0.058 (0.49)	-0.033 (0.30)
No industry given	-0.493 (1.43)	0.074 (0.32)	-0.537 (1.66)	-0.109 (0.60)
Unskilled	-0.092 (1.48)	-0.151 (2.01)	-0.057 (1.10)	-0.120 (2.57)
Skilled Manual	-0.161 (2.65)	-0.117 (1.48)	-0.109 (2.14)	-0.044 (0.90)
Professional/Intermediate	0.061 (0.69)	0.158 (1.61)	0.152 (2.09)	0.048 (0.85)
No occupation given	0.322 (1.89)	0.108 (0.69)	0.443 (2.99)	0.108 (1.06)
Youth Training				
YT One (< 12 months)	0.089 (1.67)		0.094 (2.17)	..
YT Three (> 24 months)	-0.086 (1.68)		-0.027 (0.61)	..
YT Employer	-0.049 (1.07)		-0.072 (1.90)	..
YT Qualification	0.078 (1.59)		0.073 (1.79)	..
Still on YT 1993	0.090 (1.34)		-0.038 (0.71)	..
Lambda	-0.207 (1.02)	0.010 (0.05)	-0.062 (0.54)	-0.146 (1.24)
Constant	6.594 (26.37)	6.730 (24.40)	6.424 (31.08)	6.478 (39.77)
F- statistic	2.25	3.53	3.20	5.41
Mean of dependent variable	5.9107	5.9829	5.9178	5.9955
Standard deviation of dependent	0.318	0.354	0.318	0.309
Standard error of regression	0.279	0.294	0.277	0.316
N	*218*	*180*	*308*	*433*

**coefficient (t-statistic)*

Table A8: Treatment Effects Model: wage equations for Christmas 1990 (S4) and Christmas 1991 (S5) school-leavers in employment, Autumn 1993

Descriptive variables	*S4 Leavers*	*S4 & S5 Leavers*
Personal characteristics		
Female	-0.212 (5.51)	-0.160 (5.65)
Marital status	0.016 (0.28)	0.038 (0.92)
School experience		
Unqualified	0.004 (0.08)	-0.046 (0.92)
1-3 O/S-Grades	0.045 (0.87)	0.049 (1.35)
4+ O/S-Grades	0.020 (0.27)	0.078 (1.99)
Highers	..	0.120 (2.68)
S4 Leaver	..	0.048 (1.69)
Labour market experience		
Local unemployment rate	-0.006 (1.11)	-0.006 (1.58)
Unemployment duration (mths)	-0.004 (1.05)	-0.008 (2.77)
Hours worked	-0.012 (5.66)	-0.011 (7.42)
Does Training	0.059 (1.69)	0.029 (1.06)
Self Employed	0.100 (0.92)	0.220 (2.26)
Industrial and occupational descriptors		
Mining and utilities	0.389 (2.49)	0.335 (2.72)
Manufacturing	0.130 (1.32)	0.047 (0.57)
Construction	0.003 (0.03)	-0.087 (1.03)
Services	-0.018 (0.19)	-0.059 (0.73)
No industry given	-0.066 (0.36)	-0.215 (1.41)
Unskilled	-0.122 (2.57)	-0.095 (2.75)
Skilled Manual	-0.145 (3.04)	-0.075 (2.12)
Professional/Intermediate occupation	0.099 (1.52)	0.080 (1.77)
No occupation given	0.228 (1.99)	0.209 (2.50)
Youth Training		
YT experience	-0.265 (1.02)	0.036 (0.26)
YT One (< 12 months)	0.121 (2.22)	0.111 (2.39)
YT Three (> 24 months)	-0.082 (1.56)	-0.044 (0.96)
YT Employer	-0.038 (0.89)	-0.064 (1.74)
YT Qualification	0.092 (1.85)	0.081 (1.87)
Still on YT 1993	-0.092 (1.34)	-0.008 (0.14)
Lambda	-0.087 (0.55)	-0.097 (1.21)
Constant	6.6834 (56.44)	6.485 (54.35)
F- statistic	4.57	6.67
Mean of dependent variable	5.9434	5.9632
Standard deviation of dependent variable	0.336	0.339
Standard error of regression	0.292	0.302
N	*398*	*741*

**coefficient (t-statistic)*

Table A9: Logit/OLS (Lee Model): wage equations for Christmas 1999 (S4) and Christmas 1991 (S4 & S5) school-leavers in employment, Autumn 1993

	S4 Leavers		*S4 & S5 Leavers*	
Descriptive variables	*YT*	*Non-YT*	*YT*	*Non-YT*
Personal characteristics				
Female	-0.139 (2.81)	-0.315 (5.06)	-0.094 (2.38)	-0.209 (5.32)
Marital status	0.046 (0.63)	-0.014 (0.17)	0.092 (1.52)	-0.011 (0.19)
School experience				
Unqualified	0.061 (0.88)	0.011 (0.13)	-0.013 (0.24)	-0.036 (0.50)
1-3 O/S-Grades	0.096 (1.47)	0.064 (0.81)	0.034 (0.74)	0.090 (1.65)
4+ O/S-Grades	0.170 (1.54)	0.018 (0.16)	0.066 (1.19)	0.125 (2.13)
Highers	..	..	0.118 (1.63)	0.143 (2.38)
S4 Leaver	..	..	0.042 (0.96)	0.043 (1.16)
Labour market experience				
Local unemployment rate	-0.014 (1.64)	-0.009 (0.99)	-0.002 (0.35)	-0.011 (1.94)
Unemployment duration (mths)	-0.003 (0.59)	-0.001 (0.20)	-0.007 (1.72)	-0.007 (1.83)
Hours worked	-0.011 (3.66)	-0.013 (4.57)	-0.012 (4.70)	-0.010 (5.68)
Does Training	0.060 (1.31)	0.052 (0.97)	0.066 (1.73)	0.010 (0.26)
Self Employed	-0.018 (0.12)	0.199 (1.21)	0.021 (0.16)	0.400 (2.84)
Industrial and occupational				
Mining and utilities	-0.094 (0.30)	0.556 (2.81)	-0.048 (0.15)	0.368 (2.52)
Manufacturing	0.136 (1.00)	0.157 (1.07)	0.091 (0.76)	0.042 (0.39)
Construction	0.059 (0.42)	-0.029 (0.20)	0.031 (0.25)	-0.147 (1.28)
Services	-0.021 (0.16)	0.014 (0.10)	-0.058 (0.49)	-0.033 (0.30)
No industry given	-0.493 (1.44)	0.074 (0.32)	-0.537 (1.66)	-0.109 (0.59)
Unskilled	-0.092 (1.48)	-0.151 (2.01)	-0.057 (1.11)	-0.119 (2.58)
Skilled Manual	-0.161 (2.65)	-0.117 (1.48)	-0.109 (2.14)	-0.044 (0.91)
Professional/Intermediate	0.061 (0.69)	0.158 (1.61)	0.152 (2.09)	0.048 (0.85)
No occupation given	0.322 (1.88)	0.108 (0.69)	0.443 (2.99)	0.108 (1.05)
Youth Training				
YT One (< 12 months)	0.089 (1.67)		0.094 (2.17)	..
YT Three (> 24 months)	-0.086 (1.68)		-0.027 (0.61)	..
YT Employer	-0.049 (1.07)		-0.073 (1.91)	..
YT Qualification	0.078 (1.59)		0.073 (1.79)	..
Still on YT 1993	0.090 (1.33)		-0.038 (0.71)	..
Lambda	-0.204 (1.02)	0.008 (0.04)	-0.065 (0.57)	-0.146 (1.24)
Constant	6.591 (26.55)	6.729 (24.45)	6.427 (31.12)	6.479 (39.76)
F- statistic	2.25	3.53	3.20	5.41
Mean of dependent variable	5.9107	5.9829	5.9178	5.9955
Standard deviation of dependent	0.318	0.354	0.318	0.351
Standard error of regression	0.279	0.294	0.277	0.309
N	*218*	*180*	*308*	*433*

**coefficient (t-statistic)*

Table A10: Endogenous Switching Regression: wage equations for Christmas 1999 (S4) and Christmas 1991 (S4 & S5) school-leavers in employment, Autumn 1993

Descriptive variables	*S4 Leavers* *YT*	*Non-YT*	*S4 & S5 Leavers* *YT*	*Non-YT*
Personal characteristics				
Female	-0.132 (2.44)	-0.298 (4.13)	-0.092 (2.26)	-0.216 (5.01)
Marital status	0.042 (0.56)	-0.042 (0.30)	0.094 (1.37)	-0.009 (0.13)
School experience				
Unqualified	0.064 (0.93)	-0.047 (0.49)	-0.026 (0.41)	-0.030 (0.37)
1-3 O/S-Grades	0.106 (1.71)	-0.015 (0.18)	0.032 (0.65)	0.094 (1.68)
4+ O/S-Grades	0.178 (2.07)	0.110 (1.22)	0.064 (1.08)	0.114 (2.02)
Highers	..	..	0.118 (1.42)	0.139 (2.10)
S4 Leaver	..	..	0.044 (0.88)	-0.044 (1.19)
Labour market experience				
Local unemployment rate	-0.013 (1.90)	-0.002 (0.21)	-0.003 (0.49)	-0.008 (1.44)
Unemployment duration (mths)	-0.003 (0.49)	-0.001 (0.25)	-0.008 (1.71)	-0.007 (2.32)
Hours worked	-0.010 (3.66)	-0.014 (3.92)	-0.012 (4.69)	-0.010 (5.96)
Does Training	0.060 (1.18)	0.040 (0.61)	0.067 (1.63)	0.009 (0.24)
Self Employed	-0.040 (0.16)	0.275 (1.87)	0.020 (0.09)	0.392 (4.17)
Industrial and occupational				
Mining and utilities	-0.149 (0.00)	0.493 (1.97)	-0.034 (0.00)	0.374 (2.56)
Manufacturing	0.104 (0.70)	0.124 (0.89)	0.095 (0.78)	0.048 (0.47)
Construction	0.026 (0.17)	-0.143 (0.95)	0.036 (0.29)	-0.145 (1.35)
Services	-0.048 (0.34)	0.041 (0.28)	-0.053 (0.46)	-0.030 (0.31)
No industry given	-0.556 (0.00)	0.004 (0.02)	-0.536 (0.00)	-0.102 (0.55)
Unskilled	-0.098 (1.40)	-0.163 (1.94)	-0.060 (0.97)	-0.122 (2.34)
Skilled Manual	-0.150 (2.42)	-0.110 (1.19)	-0.109 (2.12)	-0.049 (1.00)
Professional/Intermediate	0.066 (0.69)	0.151 (1.16)	0.153 (2.02)	0.049 (0.67)
No occupation given	0.325 (1.88)	0.076 (0.33)	0.440 (3.07)	0.108 (0.86)
Youth Training				
YT One (< 12 months)	0.092 (1.67)		0.093 (1.88)	..
YT Three (> 24 months)	-0.076 (1.68)		-0.026 (0.56)	..
YT Employer	-0.051 (1.07)		-0.073 (1.75)	..
YT Qualification	0.072 (1.59)		0.073 (1.76)	..
Still on YT 1993	0.108 (1.33)		-0.039 (0.79)	..
Constant	6.627 (30.18)	7.048 (27.29)	6.458 (30.43)	6.5 04 (40.88)
Log-likelihood	-324.28		-624.74	
Variance parameters				
Sigma (0)	0.403 (9.37)		0.314 (15.77)	
Rho (0, u)	-0.858 (10.73)		0.237 (0.641)	
Sigma (1)	0.345 (9.19)		0.288 (11.00)	
Rho (1, u)	-0.769 (5.95)		-0.341 (0.96)	
N	*218 (398)*	*180 (398)*	*308 (741)*	*433 (741)*

**coefficient (t-statistic)*

Table A11: Probability that a school leaver is in employment in October 1993, using SYPS data and Heteroskedastic Model*

Descriptive variables	*Whole Sample*	*S4 Leavers*
Personal characteristics		
Female	-0.021 (0.23)	-0.037 (0.20)
Married	0.296 (1.77)	0.635 (1.83)
School experience		
Unqualified	-0.095 (0.66)	-0.196 (0.77)
1-3 O/S-levels	0.157 (1.14)	0.028 (0.18)
4+ O/S-levels	0.349 (1.98)	0.451 (1.10)
Highers	0.400 (1.84)	..
S4 leaver	0.473 (2.16)	..
S5 leaver	0.193 (0.45)	..
Habitual truant	-0.299 (1.82)	-0.507 (2.16)
Part time job while at school	0.149 (1.63)	-0.042 (0.23)
Background characteristics		
Father unemployed	-0.300 (1.69)	-0.483 (1.47)
Father in skilled/ manual occupation	0.262 (2.23)	0.440 (2.17)
Father in professional/ intermediate	0.216 (1.40)	0.535 (1.34)
Father's occupation not given	0.237 (1.77)	0.447 (1.65)
Mother in non-manual occupation	0.104 (1.06)	0.210 (0.92)
Labour market experience		
Local unemployment rate	-0.016 (1.10)	0.001 (0.03)
Unemployment experience (mths)	-0.091 (8.15)	-0.091 (5.15)
Unemployed October 1991	0.364 (0.93)	0.163 (0.32)
Unemployed October 1992	-0.640 (2.91)	-1.159 (3.04)
Youth Training		
YT Experience	0.146 (0.17)	0.236 (0.54)
YT One (< 12 months)	-0.028 (0.16)	-0.271 (0.92)
YT Three (> 24 months)	0.035 (0.19)	0.108 (0.38)
YT Number of courses	-0.210 (2.01)	-0.343 (2.07)
YT Qualification	-0.010 (0.72)	0.050 (0.22)
Still on YT 1993	-0.396 (1.90)	0.107 (0.08)
Variance Function		
S5 leaver	-0.079 (1.32)	..
Unemployed October 1991	0.395 (1.25)	0.577 (1.42)
Unemployed October 1992	0.474 (1.90)	0.364 (1.26)
YT Experience	-0.058 (0.56)	-0.108 (0.50)
Still on YT 1993	-0.195 (1.90)	1.013 (1.00)
Constant	0.836 (2.83)	1.318 (2.91)
Chi-square	386.86	258.67
Log-likelihood	-432.23	-192.98
LR test for heteroskedasticity	12.77 (5 d.f.)	8.18 (4 d.f.)
LR test for more general heteroskedasticity	21.29 (16 d.f.)	17.04 (13 d.f.)
N	*1,293*	*596*

**coefficient (t-statistic)*

Appendix B: Data Description: The Australian Youth Survey

The Australian Youth Survey (AYS) is a longitudinal survey programme collected on behalf of the Department of Employment, Education, Training and Youth Affairs (DETYA)[175] and maintained by the Social Sciences Data Archives at the Australian National University, Canberra.[176] The survey began in 1989 with a nationally representative sample of individuals aged between 16 and 19. This group has been re-interviewed each year since with a new group of 16-year-olds added each year from 1990 to 1994.

The key aim of the AYS is to provide data on the dynamics of the youth labour market and enable questions to be addressed, which are not otherwise readily covered with available cross-sectional data. For example, the distinct changes in the youth labour market with respect to increased school staying on rates and the school-to-work transition. Labour market topics include detailed work history, job training, job search behaviour, and experience with the Commonwealth Employment Service (CES). Other topics related to the main labour market theme include secondary schooling, further and higher education, career advice, qualifications, school-to-work transition, health, housing and financial conditions. The survey also includes basic demographic variables such as age, gender, marital status, racial origin, parental education and occupation and area of residence.

At the outset the sample comprised 700 census collection districts covering all of Australia excluding sparsely populated areas. Ten persons aged 15 to 19 years were selected from each district giving a target sample of 7,000. Those aged 16 to 19 comprised the 1989 sample with the 15 year olds being interviewed one year later in 1990. Samples of 16 year olds added in 1991 to 1994 were identified from school-based samples of 14 year olds, two years prior to their AYS interview. A sample of 56 schools was selected for each year with probability proportional to the number of 14-year-old students, from which 25 were chosen for the subsequent AYS interview

---

[175] Formally the Commonwealth Department of Employment, Education and Training.

[176] This section draws on the technical reports and survey questionnaires that are provided with the dataset together with documentation available on the Social Sciences Data Archives website: http://ssda.anu.edu.au/.

giving a target sample of 1400 16 year olds each year. The survey instrument was personal interview carried out by an independent research group, Reark Research Pty, Ltd. The achieved sample sizes for each year of the survey, including the 16-year-olds added annually from 1990 to 1994, are given in table 7.1.

Table A12: AYS sample sizes including annual addition of 16 year olds by age group and year of collection

Year	*1989*	*1990*	*1991*	*1992*	*1993*	*1994*
Retained sample	5,350	5,350	6,250	6,947	7,607	8,021
Additional 16 year olds		1,501	1,146	1,198	1,088	1,116
Not interviewed		(601)	(449)	(538)	(674)	(787)
Total	**5,350**	**6,250**	**6,947**	**7,607**	**8,021**	**8,350**
Age group	16-19	16-20	16-21	16-22	16-23	16-24

Source: DETYA, AYS Technical Reports

The dataset contains an overall total of 11, 425 cases (un-weighted) with a potential 1253 variables per subject, in each of the six sweeps that make up the survey. However, the interviewers employ the 'skip' technique meaning that subjects were only asked questions that were relevant to them keeping the actual data-set somewhat smaller. Where appropriate interviewers inspected wage slips to validate responses to questions on wages and hours worked.

Similar to the major UK government sponsored surveys, the Scottish Young Person's Survey (SYPS) and the Youth Cohort Study (YCS), the Australian survey required respondents to complete a retrospective diary (calendar). However, personal interviewing facilitated the collection of greater detail, with respondents requested to record their major activities on a weekly basis over the 12 months preceding the interview. Hence the AYS is more comprehensive then either of the UK studies providing more data points over a longer period of time, however, attrition rates are high at more than 10 percent from year to year. Miller (1994, p. 126) claims that Australian National University studies have shown this should not be problematic if researchers employ the weightings provided.

Each year of the sample is weighted to match Australian Bureau of Statistics population benchmarks, by region (typically by State capital city and remainder of State), age and sex. Within the dataset an individual weighting variable is supplied for each respondent in each wave, to enable estimates from the sample to be compared with the known population. To help maintain uniformity with other surveys the AYS survey data, where possible, adhere to Australian Bureau of Statistics' definitions.

The school-to-work transition is detailed in the survey but very few of the respondents are identified as being on traineeships or other government assisted training programmes. Those respondents reporting that their training was part of a traineeship made up just 1.55 percent of the sample, with the majority of those joining only in the final sweep of the survey. Those individuals that indicated their training was part of Jobstart, Jobskills or other programme made up a further 3 percent of the sample, again predominately in the final sweep. Dearden (1995) combined these categories as a single descriptive in her analysis of firm training and reported a marginally positive but insignificant impact on the post-programme earnings of participants.